COVER SHEET

5972.US.P7

Filed: October 11, 2001

TITLE:

REAGENTS AND METHODS USEFUL FOR DETECTING DISEASES OF THE BREAST

The inventors of the attached Patent Application are:

Jerry Henslee
Paula N. Friedman

Submitted by:

Mimi C. Goller, Counsel
Abbott Laboratories
Dept. 377, AP6D/2
100 Abbott Park Road
Abbott Park, IL. 60064-6050
(847) 935-7550

REAGENTS AND METHODS USEFUL FOR DETECTING DISEASES OF THE BREAST

BACKGROUND OF THE INVENTION

Cross-Reference to Related Applications

This application is a continuation-in-part of pending U.S. patent application Serial No. 09/467,602 filed on December 20, 1999, which is a a continuation-in-part of U.S. patent application Serial No. 09/215,818 filed on December 18, 1998, now allowed, which is a continuation-in-part of issued U.S. patent application Serial No. 08/912,276, filed on August 15, 1997, which is a continuation-in-part of U.S. patent application Serial No. 08/697,105, filed on August 19, 1996, now abandoned, as well as a continuation-in-part of U.S. patent application Serial No. 08/912,149, filed on August 15, 1997, now abandoned, which is a continuation-in-part of U.S. patent application Serial No. 08/697,106, filed on August 19, 1996, now abandoned, and Serial No. 08/962,094 filed on October 31, 1997 and Serial No. 09/516,444 filed on February 20, 2000 from which priority is claimed pursuant to 35 U.S.C. §120 and which are all incorporated herein by reference in their entirety.

Background Information

This invention relates generally to detecting diseases of the breast. Furthermore, the invention also relates to reagents and methods for detecting diseases of the breast. More particularly, the present invention relates to reagents such as polypeptide sequences, as well as methods which utilize these sequences. The polypeptide sequences are useful for detecting, diagnosing, staging, monitoring, prognosticating, in vivo imaging, preventing or treating, or determining predisposition to diseases or conditions of the breast, such as breast cancer.

Breast cancer is the most common form of cancer occurring in females in the U.S. The incidence of breast cancers in the United States is projected to be 180,300 cases diagnosed and 43,900 breast cancer-related deaths to occur during 1998 (American Cancer Society statistics). Worldwide, the incidence of breast cancer increased from 700,000 in 1985 to about 900,000 in 1990. G.N. Hortobagyi et al., CA Cancer J Clin 45:199-226 (1995).

Procedures used for detecting, diagnosing, staging, monitoring, prognosticating, in vivo imaging, preventing or treating, or determining predisposition to diseases or conditions of the breast, such as breast cancer, are of critical importance to the outcome of the patient. For example, patients diagnosed with early breast cancer have greater than a 90% five-year relative survival rate as compared to a survival rate of about 20% for patients diagnosed with distantly metastasized breast cancers. (American Cancer Society statistics). Currently, the best initial indicators of early breast cancer are physical examination of the breast and mammography. J.R. Harris et al. In: Cancer: Principles and Practice of Oncology, Fourth Edition, pp. 1264-1332, Philadelphia, PA: J/B. Lippincott Co. (1993). Mammography may detect a breast tumor before it can be detected by physical examination, but it has limitations. For example, mammography's predictive value depends on the observer's skill and the quality of the mammogram. In addition, 80 to 93% of suspicious mammograms are false positives, and 10 to 15% of women with breast cancer have false negative mammograms. C.J. Wright et al., Lancet 346:29-32 (1995). New diagnostic methods which are more sensitive and specific for detecting early breast cancer are clearly needed.

Breast cancer patients are closely monitored following initial therapy and during adjuvant therapy to determine response to therapy, and to detect persistent or recurrent disease, or early distant metastasis. Current diagnostic procedures for monitoring breast cancer include mammography, bone scan, chest radiographs, liver function tests and tests for serum markers. The serum tumor markers most commonly used for monitoring patients are carcinoembryonic antigen (CEA) and CA 15-3. Limitations of CEA include absence of elevated serum levels in about 40% of women with metastatic disease. In addition, CEA elevation during adjuvant therapy may not be related to recurrence but to other factors that are not clinically important. CA 15-3 can also be negative in a significant number of patients with progressive disease and, therefore, fail to predict metastasis. Both CEA and CA 15-3 can be elevated in nonmalignant, benign conditions giving rise to false positive results. Therefore, it would be clinically beneficial to find a breast associated marker which is more sensitive and specific in detecting cancer recurrence. J. R. Harris et al., supra. M. K. Schwartz, In: Cancer: Principles and Practice of Oncology, Vol. 1, Fourth Edition, pp. 531 - 542, Philadelphia, PA: J/B.

Lippincott Co. 1993.

Another important step in managing breast cancer is to determine the stage of the patient's disease because stage determination has potential prognostic value and provides criteria for designing optimal therapy. Currently, pathological staging of breast cancer is preferable over clinical staging because the former gives a more accurate prognosis. J. R. Harris et al., supra. On the other hand, clinical staging would be preferred were it at least as accurate as pathological staging because it does not depend on an invasive procedure to obtain tissue for pathological evaluation. Staging of breast cancer could be improved by detecting new markers in serum or urine which could differentiate between different stages of invasion. Such markers could be protein markers expressed by cells originating from the primary tumor in the breast but residing in blood, bone marrow or lymph nodes and could serve as sensitive indicators for metastasis to these distal organs. For example, specific protein antigens, associated with breast epithelial cells, have been detected by immunohistochemical techniques, in bone marrow, lymph nodes and blood of breast cancer patients suggesting metastasis. K. Pantel et al., Onkologie 18:394-401 (1995).

Such diagnostic procedures also could include immunological assays based upon the appearance of various disease markers in test samples such as blood, plasma, serum or urine obtained by minimally invasive procedures which are detectable by immunological methods. These diagnostic procedures would provide information to aid the physician in managing the patient with disease of the breast, at low cost to the patient. Markers such as prostate specific antigen (PSA) and human chorionic gonadotropin (hCG) exist and are used clinically for screening patients for prostate cancer and testicular cancer, respectively. For example, PSA normally is secreted by the prostate at high levels into the seminal fluid, but is present in very low levels in the blood of men with normal prostates. Elevated levels of PSA protein in serum are used in the early detection of prostate cancer or disease in asymptomatic men. See, for example, G.E. Hanks et al., In: Cancer: Principles and Practice of Oncology, Vol. 1, Fourth Edition, pp. 1073-1113, Philadelphia, PA: J.B. Lippincott Co. 1993. M. K. Schwartz et al., In: Cancer: Principles and Practice of Oncology, Vol. 1, Fourth Edition, pp. 531-542, Philadelphia, PA: J.B. Lippincott Co. 1993. Likewise, the management of breast diseases could be improved by the use of new markers normally expressed in the breast but found in elevated amounts in an inappropriate

body compartment as a result of the disease of the breast.

Further, new markers which could predict the biologic behavior of early breast cancers would also be of significant value. Early breast cancers that threaten or will threaten the life of the patient are more clinically important than those that do not or will not be a threat. G.E. Hanks, supra. Such markers are needed to predict which patients with histologically negative lymph nodes will experience recurrence of cancer and also to predict which cases of ductal carcinoma in situ will develop into invasive breast carcinoma. More accurate prognostic markers would allow the clinician to accurately identify early cancers localized to the breast which will progress and metastasize if not treated aggressively. Additionally, the absence of a marker for an aggressive cancer in the patient could spare the patient expensive and non-beneficial treatment. J. R. Harris et al., supra. E. R. Frykberg et al., Cancer 74:350-361 (1994).

It therefore would be advantageous to provide specific methods and reagents useful for detecting, diagnosing, staging, monitoring, prognosticating, in vivo imaging, preventing or treating, or determining predisposition to diseases or conditions of the breast. Such methods would include assaying a test sample for products of a gene which are overexpressed in diseases and conditions associated with the breast, including cancer. Such methods may further include assaying a test sample for products of a gene whose distribution among the various tissues and compartments of the body have been altered by a breast-associated disease or condition, including cancer. Such methods would comprise making cDNA from mRNA in the test sample, amplifying, when necessary, portions of the cDNA corresponding to the gene or a fragment thereof, and detecting the cDNA product as an indication of the presence of the disease or condition including cancer or detecting translation products of the mRNAs comprising gene sequences as an indication of the presence of the disease. Useful reagents include polynucleotide(s), or fragment(s) thereof which may be used in diagnostic methods such as reverse transcriptase-polymerase chain reaction (RT-PCR), PCR, or hybridization assays of mRNA extracted from biopsied tissue, blood or other test samples; or proteins which are the translation products of such mRNAs; or antibodies directed against these proteins. Such assays would include methods for assaying a sample for product(s) of the gene and detecting the product(s) as an indication of disease of the breast. For example, these assays would include methods for detecting the gene products (proteins) in light of possible post-translational

modifications that can occur in the body. Such post-translational modifications can include proteolytic processing, alteration of the chain termini, glycosylation, lipid attachment, sulfation, gamma-carboxylation, hydroxylation, phosphorylation, ADP-ribosylation, disulfide bond formation, and multiple non-covalent interactions with molecules such as co-factors, inhibitors (both small molecule and protein), activators (both small molecule and protein), and other proteins in formation of multi-subunit complexes. See, for example, T. E. Creighton et al., In: Proteins: Structures and Molecular Properties, Second Edition, pp. 78-102, New York, NY:W. H. Freeman and Co. 1993. Some modifications are sequence specific and are therefore predictive whereas others are not and are observed by empirical data only.

The uteroglobin family of proteins contains a small number of sequences whose function has yet to be identified but may serve in detecting, diagnosing, staging, monitoring, prognosticating, in vivo imaging, preventing or treating, or determining predisposition to diseases or conditions of the breast. L. Miele et al., J Endocrinol. Invest. 17:679-692 (1994). Empirically, uteroglobins have been found to complex with another molecule of themselves, forming a homo-dimeric multi-subunit complex. R.Bally et al., J. Mol Biol 206:153-170(1989); I. Morize et al., J. Mol Biol 194:725-739 (1987); T. C. Umland et al., Nature Structural Biology 1:538-545 (1994); T. C. Umland et al., J. Mol. Biol. 224:441-448(1992). Other sequences that appear to be distantly related to uteroglobins include the rat steroid binding protein and the cat major allergen. Like the uteroglobins, these proteins have been determined to exist as multi-subunit complexes. Unlike uteroglobins, these subunits are heterodimeric, i.e., from different sequences. Furthermore, these heterodimers complex together with either another copy of themselves forming an $\alpha\beta/\alpha\beta$ heterotetramer, as in the cat major allergen [O. A.Duffort et al., Molecular Immunology 28:301-309 (1991); K. Leitermann et al., J of Allergy and Clinical Immunology 74:147-153 (1991)], or they complex with a different heterodimer, such as the rat steroid binding protein which has the subunit structure $\alpha\beta/\alpha'\beta$ (where α and α' are homologous but not identical). M. Parker et al., Nature 298:92-94 (1982). In the case of the cat major allergen, α is homologous to the uteroglobin family but β is not. J. P. Morgenstern et al., PNAS 88:9690-9694(1991). In the case of the rat steroid binding protein, α, α', and β have varying degrees of homology to the uteroglobin family of proteins.

Mammaglobin has recently been described as a newly discovered addition to the uteroglobin family, albeit a distantly related member. Its expression is reported to be restricted to mammary epithelium by Northern blot and RT/PCR analysis. M. A. Watson et al., Cancer Research 56:860-865 (1996). The gene has been localized to chromosome 11q13, and several potential transcriptional control elements have been identified. M. A. Watson et al., Oncogene 16:817-824 (1998). Furthermore, the polynucleotide sequence was described in U.S. Patent 5,668,267. However, there are no reports describing the nature of the protein product.

BU101 was first described as an endometrial specific uteroglobin (WO 97/34997). In contrast, the present inventors have recently described BU101 as a breast specific uteroglobin (see U.S. patent application serial number 08/697,105 filed on 8/19/96 which was abandoned in favor of continuation-in-part U.S. patent application serial number 08\912,276 filed on 8/15/97). Its detection in breast clinical specimens was shown in these previous applications. The nature of Mammaglobin and BU101 protein products is newly described in this application.

Drug treatment or gene therapy for diseases and conditions of the breast including cancer can be based on these identified gene sequences or their expressed proteins, and efficacy of any particular therapy can be monitored. Furthermore, it would be advantageous to have available alternative, non-surgical diagnostic methods capable of detecting early stage breast disease, such as cancer.

All U.S. patents and publications referred to herein are hereby incorporated in their entirety by reference.

SUMMARY OF THE INVENTION

The present invention provides a new entity, specifically, a multimeric polypeptide complex, wherein at least one copy of BU101 polypeptide (SEQUENCE ID NO 6) and at least one copy of Mammaglobin polypeptide (SEQUENCE ID NO 5) are present but may contain one or more unknown polypeptides as well. Mammaglobin polypeptide may be present as a glycoprotein, with sugars attached at asparagine residues located at position 53, and/or position 68, or neither. Furthermore, Mammaglobin polypeptide may be linked covalently via disulfide bonds to BU101 polypeptide. Both sequences contain 3 cysteine residues in their mature form. This disulfide linked heterodimer may constitute one subunit of the complex and it may have interactactions with another subunit of identical composition, forming

an $\alpha\beta/\alpha\beta$ heterotetramer; or it may interact with a subunit of nonidentical composition, forming an $\alpha\beta/\alpha'\beta$, or an $\alpha\beta/\alpha\beta'$, or an $\alpha\beta/\alpha'\beta'$ heterotetramer, where β represents BU101 polypeptide, β represents Mammaglobin polypeptide, α' represents a polypeptide homologous to but not identical to BU101 polypeptide, and β' represents a polypeptide homologous to but not identical to Mammaglobin polypeptide. The gene encoding the BU101 polypeptide may contain a single base T/C polymorphism which results in either a proline residue (encoded by CCG) or a leucine residue (encoded by CTG) at amino acid 53 of the polypeptide. The multimeric polypeptide complex can be produced by recombinant technology, produced by isolation from natural sources, or produced by synthetic techniques.

A method for producing a polypeptide, or polypeptide complex, which contains at least one epitope of a multimeric polypeptide complex is provided, which method comprises incubating host cells transfected with one or more expression vector(s). The vector(s) comprises a polynucleotide sequence encoding one or more polypeptide(s), wherein the polypeptide(s) comprises an amino acid sequence having at least 50% identity with an amino acid sequence selected from the group consisting of BU101 polypeptide (SEQUENCE ID NO 6), Mammaglobin polypeptide (SEQUENCE ID NO 5), or an unknown α' or β' polypeptide, and fragments thereof.

The present invention provides a cell co-transfected with nucleic acid sequences that encode at least one component polypeptide sequence of a multimeric polypeptide antigen, or fragments thereof. The nucleic acid sequence is selected from the group consisting of BU101 (SEQUENCE ID NO 2), Mammaglobin (SEQUENCE ID NO 1), α', or β', and fragments or complements thereof.

A method for producing antibodies to antigens consisting of either BU101 polypeptide (SEQUENCE ID NO 6), Mammaglobin polypeptide (SEQUENCE ID NO 5), unknown α' or β' polypeptide, or a multimeric polypeptide complex, or fragments thereof, also is provided, which method comprises administering to an individual an isolated immunogenic polypeptide, polypeptide complex, or fragment thereof, wherein the isolated immunogenic polypeptide comprises at least one epitope of the multimeric polypeptide complex, wherein the at least one MPA epitope has at least 20% identity to an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:6, and fragments thereof.

The immunogenic polypeptide, polypeptide complex, or fragment thereof is administered in an amount sufficient to produce an immune response. The isolated,

immunogenic polypeptide, polypeptide complex, or fragment thereof comprises an amino acid sequence selected from the group consisting of BU101 polypeptide (SEQUENCE ID NO 6), Mammaglobin polypeptide (SEQUENCE ID NO 5), unknown α', or β' polypeptide sequence, and fragments thereof.

Also provided is an antibody which specifically binds to at least one epitope of the multimeric polypeptide complex. The antibody can be a polyclonal or monoclonal antibody. The epitope is derived from an amino acid sequence selected from the group consisting of BU101 polypeptide (SEQUENCE ID NO 6), Mammaglobin polypeptide (SEQUENCE ID NO 5), unknown α' or β' polypeptides, or any combination thereof. That is, the epitope may be shared between polypeptide sequences.

Also provided is an antibody which specifically binds to at least one epitope of the BU101 polypeptide (SEQUENCE ID NO 6). The antibody may or may not bind to the multimeric polypeptide complex. The antibody can be a polyclonal or monoclonal antibody. The epitope is derived from an amino acid sequence selected from the group consisting of BU101 polypeptide (SEQUENCE ID NO 6), and fragments thereof.

Also provided is an antibody which specifically binds to at least one epitope of the Mammaglobin polypeptide (SEQUENCE ID NO 5). The antibody may or may not bind to the multimeric polypeptide complex. The antibody can be a polyclonal or monoclonal antibody. The epitope is derived from an amino acid sequence selected from the group consisting of Mammaglobin polypeptide (SEQUENCE ID NO 5), and fragments thereof.

Also provided is an antibody which specifically binds to at least one epitope of an unknown α' polypeptide. The antibody may or may not bind to the multimeric polypeptide complex. The antibody can be a polyclonal or monoclonal antibody.

Also provided is an antibody which specifically binds to at least one epitope of an unknown β' polypeptide. The antibody may or may not bind to the multimeric polypeptide complex. The antibody can be a polyclonal or monoclonal antibody.

A method for detecting the multimeric polypeptide antigen in a test sample suspected of containing the multimeric polypeptide antigen also is provided. The method comprises contacting the test sample with an antibody or fragment thereof which specifically binds to at least one epitope of the multimeric polypeptide antigen, for a time and under conditions sufficient for the formation of antibody/antigen

complexes; and detecting the presence of such complexes containing the antibody as an indication of the presence of the multimeric polypeptide antigen in the test sample. The antibody can be attached to a solid phase and may be either a monoclonal or polyclonal antibody.

Assay kits for determining the presence of the multimeric polypeptide antigen in a test sample are also included. In one embodiment, the assay kit comprises a container containing an antibody which specifically binds to a multimeric polypeptide antigen, wherein the antigen comprises at least one epitope encoded by either the BU101 gene, the Mammaglobin gene, the α' gene, or the β' gene. These test kits can further comprise containers with tools useful for collecting test samples (such as blood, urine, saliva, and stool). Such tools include lancets and absorbent paper or cloth for collecting and stabilizing blood; swabs for collecting and stabilizing saliva; cups for collecting and stabilizing urine or stool samples. Collection materials, such as papers, cloths, swabs, cups and the like, may optionally be treated to avoid denaturation or irreversible adsorption of the sample. These collection materials also may be treated with, or contain, preservatives, stabilizers or antimicrobial agents to help maintain the integrity of the specimens. The antibody can be attached to a solid phase.

Another method is provided which detects antibodies which specifically bind to the multimeric polypeptide antigen in a test sample suspected of containing these antibodies. The method comprises contacting the test sample with a polypeptide which contains at least one epitope of the multimeric polypeptide complex. Contacting is performed for a time and under conditions sufficient to allow antigen/antibody complexes to form. The method further entails detecting complexes which contain the polypeptide. The polypeptide complex can be attached to a solid phase. Further, the polypeptide complex can be produced recombinantly, or synthetically, or purified from natural sources.

Assay kits for determining the presence of anti-multimeric polypeptide complex antibody in a test sample or the MPA itself are also included. In one embodiment, the assay kits comprise a container containing at least one polypeptide of the multimeric polypeptide complex or the full complex itself. Further, the test kit can comprise a container with tools useful for collecting test samples (such as blood, tissue, urine, saliva, and stool). Such tools include lancets and absorbent paper or cloth for collecting and stabilizing blood; swabs for collecting and stabilizing saliva;

and cups for collecting and stabilizing urine or stool samples. Collection materials such as papers, cloths, swabs, cups, and the like, may optionally be treated to avoid denaturation or irreversible adsorption of the sample. These collection materials also may be treated with or contain preservatives, stabilizers or antimicrobial agents to help maintain the integrity of the specimens. Also, the polypeptide can be attached to a solid phase. If the antigen itself is to be detected, then an antibody is present in the container of the kit. Furthermore, if one is attempting to detect the antigen, a detergent and/or a reducing agent may also be present in the kit.

In another embodiment of the invention, antibodies or fragments thereof, against the multimeric polypeptide antigen can be used to detect or image localization of the antigen in a patient for the purpose of detecting or diagnosing a disease or condition. Such antibodies can be polyclonal or monoclonal, or made by molecular biology techniques, and can be labeled with a variety of detectable labels, including but not limited to, radioisotopes and paramagnetic metals. Furthermore, antibodies or fragments thereof, whether monoclonal, polyclonal, or made by molecular biology techniques, can be used as therapeutic agents for the treatment of diseases characterized by expression of the multimeric polypeptide antigen. In the case of therapeutic applications, the antibody may be used without derivitization, or it may be derivitized with a cytotoxic agent such as a radioisotope, enzyme, toxin, drug, prodrug, or the like.

The present invention also encompasses an additional method of detecting the presence of a multimeric polypeptide antigen (MPA) in a test sample suspected of containing the MPA. The MPA comprises at least one BU101 polypeptide and at least one Mammaglobin polypeptide. This method comprises the steps of: (a) contacting the test sample with at least one antibody specific for at least one epitope of the MPA for a time and under conditions sufficient to allow the formation of MPA/antibody complexes; adding a conjugate to said resulting MPA/antibody complexes for a time and underconditions sufficient to allow the conjugate to bind to the bound antigen, wherein the conjugate comprises an antibody attached to a signal generating compound capable of generating a detectable signal; and (c) detecting the presence of the MPA which may be present in the test sample by detecting the signal generated by the signal generating compound. The multimeric peptide antigen detected in this method and in the other methods described herein may further comprise at least one polypeptide having at least 20% identity with an amino acid sequence selected from

the group consisting of SEQ ID NO:5, SEQ ID NO:6, and fragments thereof. The antibody used in the method of step (a) is generated against a MPA, which is produced by a HEK-293 MB8 cell or a host cell transfected with a vector comprising a construct comprising at least one nucleotide sequence encoding at least one BU101 polypeptide and at least one nucleotide sequence encoding at least one Mammaglobin polypeptide. The antibodies used in the diagnostic steps of all of the present assays (or in the kits) may be produced in this manner, if so desired. Additionally, the epitope of step (a) of the above method (and of the other methods having a comparable step (a) described herein), against which the antibody is generated, may be derived from a MPA produced by a host cell comprising two vectors, wherein one of the vectors comprises a construct comprising at least one nucleotide sequence encoding at least one BU101 polypeptide and wherein the other of the two vectors comprises a construct comprising at least one nucleotide sequence encoding at least one Mammaglobin polypeptide. It should be noted that in all of the MPA detection methods of the present invention, the presence of MPA may be indicative of breast cancer or of another breast condition.

Furthermore, the present invention also includes a method of detecting the presence of a multimeric polypeptide antigen (MPA) in a test sample suspected of containing the MPA, wherein the MPA comprises at least one BU101 polypeptide and at least one Mammaglobin polypeptide, the method comprising the steps of: (a) contacting the test sample with at least one antibody specific for at least one epitope of the MPA for a time and under conditions sufficient to allow the formation of MPA/antibody complexes; (b) adding a conjugate to the resulting MPA/antibody complexes for a time and under conditions sufficient to allow the conjugate to bind to the bound antigen, wherein the conjugate comprises a steroid or antibody, attached to a signal generating compound capable of generating a detectable signal; and (c) detecting the presence of the MPA which may be present in the test sample by detecting the signal generated by the signal generating compound. The steroid may be, for example, progesterone, aldosterone, androstenedione, corticosterone, cortisol, dehydroepiandrosterone, dihydrotestosterone, estradiol, estriol, estrone, hydroxyprogesterone, and testosterone. The antibody may be generated using a MPA generated in the manner described above.

The present invention also includes a method of detecting the presence of antibody specific for a multimeric polypeptide antigen (MPA) in a test sample

suspected of containing the antibody, wherein the MPA comprises at least one BU101 polypeptide and at least one Mammaglobin polypeptide. The method comprises the steps of: (a) contacting the test sample with at least one MPA epitope derived from an amino acid sequence or fragment thereof having at least 20% identity to an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:6, and fragments thereof, for a time and under conditions sufficient to allow the formation of MPA/antibody complexes; (b) adding a conjugate to the resulting MPA/antibody complexes for a time and under conditions sufficient to allow the conjugate to bind to the bound antigen, wherein the conjugate comprises an antibody, which binds with the antibody in the test sample, attached to a signal generating compound capable of generating a detectable signal; and (c) detecting the presence of the antibody which may be present in the test sample by detecting the signal generated by the signal generating compound.

The present invention also includes a composition of matter comprising a multimeric polypeptide antigen, wherein the antigen comprises at least one BU101 polypeptide and at least one Mammaglobin polypeptide. Again, the antigen further comprises at least one polypeptide having at least 20% identity with an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:6, and fragments thereof. The composition may further comprise at least one antibody, bound to the multimeric polypeptide antigen, wherein the antibody is specific to at least one polypeptide selected from the group consisting of a BU101 polypeptide, a Mammaglobin polypeptide, one polypeptide having at least 20% identity with an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:6, and fragments thereof. In particular, two antibodies may be present and each will bind to a separate polypeptide having an amino acid sequence having at least 20% identity with an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:6, and fragments thereof. More specifically, each of the two antibodies may bind to a BU101 polypeptide or a fragment thereof, or each of the two antibodies may bind to a Mammaglobin polypeptide or a fragment thereof. Alternatively, one of the antibodies may bind to a BU101 polypeptide or a fragment thereof and the other antibody may bind to a Mammaglobin polypeptide or fragment thereof. When one of the two antibodies binds to a BU101 polypeptide or fragment thereof, the other may bind to a polypeptide having an amino acid sequence having at least 20% identity with an amino acid sequence selected from the group consisting of

SEQ ID NO:5, SEQ ID NO:6, and fragments thereof. When one of the two antibodies binds to a Mammaglobin polypeptide or fragment thereof , the other may bind to a polypeptide having an amino acid sequence having at least 20% identity with an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:6, and fragments thereof.

Moreover, the present invention also includes a method of detecting breast cancer in a patient suspected of having breast cancer comprising the steps of: (a) administering a labelled antibody specific to a multimeric protein antigen (MPA), to the patient, wherein the MPA comprises at least one BU101 polypeptide and at least one Mammaglobin polypeptide; and (b) localizing presence of the label, presence of the label indicating presence of MPA and breast cancer in the patient.

The present invention also includes method of treating breast cancer, as noted above. One such method comprises administering an antibody specific to a multimeric polypeptide antigen (MPA) to the patient, the MPA comprising at least one BU101 polypeptide and at least one Mammaglobin polypeptide.

Additionally, the present invention includes a method of diagnosing breast cancer in a patient suspecting of having breast cancer comprising the steps of: (a) preparing a tissue section or cell culture derived from a tumor excised from the patient; (b) exposing the tissue section or cell culture to an antibody specific for at least one polypeptide of a multimeric polypeptide antigen (MPA) for a time and under conditions sufficient to allow formation of antigen/antibody complexes, the polypeptide selected from the group consisting of: a BU101 polypeptide, a Mammaglobin polypeptide, a polypeptide having at least 20% identity with an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:6, and fragments thereof; and (c) localizing presence of the complexes in the tissue section or cell culture, presence of the complexes indicating presence of MPA and breast cancer in the patient.

Another method of diagnosing breast cancer in a patient suspected of having breast cancer, encompassed by the present invention, comprises the steps of detecting the presence or absence of at least one polypeptide of a multimeric polypeptide antigen (MPA), said polypeptide selected from the group consisting of a BU101 polypeptide, a Mammaglobin polypeptide, and a polypeptide having at least 20% identity with an amino acid sequence elected from the group consisting of SEQ ID NO:5, SEQ ID NO:6, and fragments thereof, in a biological sample from said patient,

presence of the at least one polypeptide indicating presence of MPA and breast cancer in the patient. The biological sample may be, for example, tissue, urine, saliva, stool, bone marrow or blood.

Another method of diagnosing breast cancer in a patient suspected of having breast cancer, and included within the present invention, comprises the steps of detecting the presence or absence of extracellular BU101 in the patient, presence of extracellular BU101 indicating breast cancer in the patient and transport of BU101 outside cells via Mammaglobin in a multimeric polypeptide antigen (MPA), said MPA comprising at least one BU101 polypeptide and at least one Mammaglobin polypeptide.

The present invention also encompasses a further method of detecting breast cancer in a patient suspected of having breast cancer comprising the steps of: (a) obtaining a biological sample from the patient; (b) measuring the amount of free BU101 polypeptide in the biological sample; (c) measuring the amount of BU101 polypeptide, present in said biological sample, complexed to Mammaglobin polypeptide; and (d) comparing the ratio of free BU101 polypeptide to complexed BU101 polypeptide, a ratio higher than 1 indicating presence of breast cancer in the patient.

The present invention also includes a method of detecting breast cancer in a patient suspecting of having breast cancer comprising the steps of: (a) obtaining a biological sample from the patient; (b) measuring the amount of free Mammaglobin polypeptide in the biological sample; (c) measuring the amount of Mammaglobin polypeptide, present in the biological sample, complexed to BU101 polypeptide; and (d) comparing the ratio of free Mammaglobin polypeptide to complexed Mammaglobin polypeptide, a ratio higher than 1 indicating presence of breast cancer in the patient.

Additionally, the present invention encompasses a method for enhancing recognition of MPA, in an immunoassay for MPA, comprising exposing the MPA (or test sample) to at least one member selected from the group consisting of a reducing agent and a detergent, prior to contacting the suspected MPA with an antibody or chemical compound. The suspected MPA may also be exposed to heat. It should be noted that in all of the methods described above and below involving detection of a MPA, the test sample or MPA may be exposed to either a reducing agent and/or a

detergent, prior to contacting the MPA with an antibody or a chemical compound. Heat may or may not also be utilized.

Furthermore, the present invention also includes a method for dissociating MPA comprising exposing said MPA to at least one member selected from the group consisting of a reducing agent and a detergent. Heat may additionally be used.

The present invention also includes a diagnostic reagent produced by an MB8 cell or a host cell transfected with a vector comprising a construct comprising at least one nucleotide sequence which encodes at least one BU101 polypeptide and at least one nucleotide sequence which encodes at least one Mammaglobin polypeptide. This reagent may be, for example, MPA.

Additionally, the present invention includes a diagnostic reagent produced by a host cell transfected with two vectors wherein one of the two vectors comprises a construct comprising at least one nucleotide sequence which encodes at least one BU101 polypeptide and wherein the other of the two vectors comprises a construct comprising at least one nucleotide sequence which encodes at least one Mammaglobin polypeptide.

Also, the present invention includes an isolated cell from the cell line HEK293-MB8 as well as the generated cell line itself.

Furthermore, the present invention also encompasses a method for detecting the presence of a multimeric polypeptide antigen (MPA) in a test sample suspected of containing the MPA, wherein the MPA comprises at least one BU101 polypeptide and at least one Mammaglobin polypeptide, the method comprising the steps of: (a) contacting the test sample with a labelled antigen selected from the group consisting of a MPA, a polypeptide of a MPA, and fragments thereof; (b) contacting the test sample and labelled antigen of step (a) with anti-MPA antibody for a time and under conditions sufficient to allow for the formation of MPA/anti-MPA antibody complexes; and (c) detecting the presence of MPA which may be present in the test sample by detecting the signal generated by the label in the labelled antigen.

Also, the present invention encompasses a method for detecting the presence of a MPA in a test sample suspected of containing the MPA, wherein the MPA comprises at least one BU101 polypeptide and at least one Mammaglobin polypeptide, the method comprising the steps of (a) contacting the test sample with a labelled antibody which binds to MPA for a time and under conditions sufficient for the formation of test sample MPA/labelled antibody complexes; (b) contacting the

step (a) complexes with an antigen selected from the group consisting of a MPA, a polypeptide of a MPA, a fragment of a MPA and a fragment of a polypeptide of a MPA, for a time and under conditions sufficient for the formation of antigen/labelled antibody complexes; and (c) detecting the presence of a signal generated by the labelled antibody wherein the signal is indicative of the presence of MPA in the test sample.

Additionally, the present invention includes a method for detecting the presence of a multimeric polypeptide antigen (MPA) in a test sample suspected of containing said MPA, wherein said MPA comprises at least one BU101 polypeptide and at least one Mammaglobin polypeptide, said method comprising the steps of: (a) contacting the test sample with a labelled steroid which binds to MPA for a time and under conditions sufficient for the formation of MPA/labelled steroid complexes; (b) contacting the MPA/labelled steroid complexes of step (a) with an antigen selected from the group consisting of a MPA, a polypeptide of a MPA, a fragment of a MPA, a fragment of a polypeptide of a MPA, for a time and under conditions sufficient to allow for the formation of antigen/labelled steroid complexes; and (c) detecting the presence of the MPA which may be present in the test sample by detecting the signal generated by the label of the labelled steroid.

The present invention also encompasses a method for detecting the presence of a MPA in a test sample suspected of containing the MPA, wherein the MPA comprises at least on BU101 polypeptide and at least one Mammaglobin polypeptide, the method comprising the steps of: (a) contacting the test sample with a steroid for a time and under conditions sufficient to allow for the formation of MPA/steroid complexes; (b) adding a conjugate to the resulting MPA/steroid complexes for a time and under conditions sufficient to allow the conjugate to bind to the bound MPA, wherein the conjugate comprises an antibody attached to a signal generating compound capable of generating a detectable signal; and (c) detecting the presence of the MPA which may be present in the test sample by detecting the signal generated by the signal generating compound.

Furthermore, the present invention also includes a method for detecting the presence of antibody specific for a multimeric polypeptide antigen (MPA) in a test sample suspected of containing the antibody, the method comprising the steps of: (a) contacting the test sample with an anti-antibody specific for the antibody for a time and under conditions sufficient to allow for the formation of antibody/anti-antibody

complexes; (b) adding a conjugate to the resulting antibody/anti-antibody complexes for a time and under conditions sufficient to allow the conjugate to bind to the bound antibody, wherein the conjugate comprises MPA attached to a signal generating compound capable of generating a detectable signal; and (c) detecting the presence of antibody which may be present in the test sample by detecting the signal generated by the signal generating compound.

Moreover, the present invention also encompasses a method for detecting the presence of an antibody specific for a MPA in a test sample suspected of containing the antibody, the method comprising the steps of (a) contacting the test sample with a labelled antigen selected from the group consisting of a MPA, a polypeptide of a MPA, a fragment of a MPA, and a fragment of a polypeptide of a MPA, for a time and under conditions sufficient to allow for the formation of antibody/labelled antigen complexes; (b) contacting the resulting complexes of step (a) with an antibody which binds to MPA, for a time and under conditions sufficient to all unbound, labelled antigen to bind to said antibody which binds to MPA; and (c) detecting the presence of the antibody which may be present in the test sample by detecting the signal generated by the labelled antigen.

Also, the present invention includes a method for detecting the presence of an antibody specific for a MPA in a test sample suspected of containing the antibody, the method comprising the steps of: (a) contacting the test sample with a MPA complexed with a steroid for a time and under conditions sufficient to allow for the formation of antibody/MPA/steroid complexes; (b) adding a conjugate to the resulting antibody/MPA/steroid complexes for a time and under conditions sufficient to allow the conjugate to bind to the bound antibody, wherein the conjugate comprises an antibody, reactive with said antibody in said test sample, attached to a signal generating compound capable of generating a detectable signal; and (c) detecting the presence of the antibody which may be present in the test sample by detecting the signal generated by the signal generating compound.

Another aspect of this invention involves a panel of at least one breast specific marker, BU101 with another breast specific marker, BS106, a newly-described mucin-like protein. BS106 is a novel mucin-like protein that was discovered through data mining of the Incyte database for breast specific markers. It was found in 310 ESTs in 64 breast tissue libraries and 5 ESTs in 1,228 libraries from other tissues.

The BS106 gene is located on chromosome 12, is composed of 4 exons and is 90 amino acids with a predicted Mr=11kd-containing a secretion signal. BS106 does not have significant homology to other genes but it does contain a mucin-like small tandem repeat-TTAAXTTA and multiple sites for O-linked glycosylation. Likewise a panel also including mammaglobin with either or both BU101 and BS106 is also contemplated.

Mammaglobin and BU101 are both members of the secretoglobin superfamily of which 15 members have so far been identified from various species. Antibodies developed against BU101 detect a specific band of 18 kd under reducing conditions and 24kd under non-reducing conditions. The identity of the larger species was shown to be a heterodimer of mammaglobin and BU101. In order for expression markers to be useful, the majority of breast cancers must express detectable levels of the genes. In the current study, the inventors have investigated the expression patterns of these two markers in a substantial panel of primary breast cancers. In addition, the inventors compared these new markers to the expression of mammaglobin and cytokeratin, two markers that have been tested for their efficacy in detecting occult breast cancer cells; (seeZach, O., Kasparu, H., Krieger, O., Hehenwarter, W., Girschikofsky, M., and Lutz, D. Detection of circulating mammary carcinoma cells in the peripheral blood of breast cancer patients via a nested reverse transcriptase polymerase chain reaction assay for mammaglobin mRNA. J. Clin. Oncol. *17*: 2015-2019, 1999; Grunewald, K., Haun, M., Urbanek, M., Fiegl, M., Muller-Holzner, E., Gunsilius, E., Dunser, M., Marth, C., and Gastl, G. Mammaglobin gene expression: a superior marker of breast cancer cells in peripheral blood in comparison to epidermal-growth-factor receptor and cytokeratin-19. Lab. Invest. *80*: 1071-1077, 2000; Kahn, H.J., Yang, L.Y., Lickley, L., Holloway, C., Hanna, W., Narod, S., McCready, D.R., Seth, A., and Marks, A. RT-PCR amplification of CK19 mRNA in the blood of breast cancer patients: correlation with established prognostic parameters. Breast Cancer Res. Treat. *60*: 143-151, 2000; Miyashiro, I., Kuo, C., Huynh, K., Iida, A., Morton, D., Bilchik, A., Giuliano, A., and Hoon, D.S. Molecular strategy of detecting metastatic cancers with use of multiple tumor-specific MAGE-A genes. Clinical Chemistry, *47*: 505-512, 2001).

The benefit from this aspect of the invention is that the markers individually have a limited expression profile; however a combination of two or more of these

markers are useful in detecting the vast majority of breast cancers. In other words, a panel of the different markers provides comprehensive analysis for the diagnosis of breast cancer as opposed to their use individually.

Further, another aspect of this invention involves assaying for the presence of BS106 and MPA. A panel with this combination also gives a more comprehensive analysis of a patient's diagnosis than using the markers individually.

BRIEF DESCRIPTION OF THE DRAWINGS

FIGURE 1 shows the binding curves of three of the monoclonal antibodies that recognized the recombinant polypeptide complex, produced in accordance with Example 7C.

FIGURE 2 shows the results of immunohistochemically staining two malignant breast sections, one normal breast section, and the HEK293-MB8 cell line with monoclonal antibody H9C65.

FIGURE 3 shows the results of immunohistochemically staining two malignant breast sections, one normal breast section, and the HEK293-MB8 cell line with monoclonal antibody J95C30.

FIGURE 4 is a scan of three Western blots showing three supernatants harvested from the growth of HEK293-MB8 cells. Blot 1 was developed with an anti-myc monoclonal antibody. Blot 2 was developed with an anti-BU101 polyclonal antisera. Blot 3 was developed with an anti-Mam polyclonal antisera.

FIGURE 5 is a scan of two dot blots showing immunorecognition of material by an anti-myc monoclonal antibody. The upper blot shows the fractions from supernatant of the MB8 cells eluting from a Nickel-chelation column. The lower blot shows the fractions from supernatant of the Mam M/H transient transfection of HEK293 cells eluting from a Nickel-chelation column.

FIGURE 6 is a scan of 4 Western blots comprising 16 panels. Supernatants from the MB8 cells and the transient transfection of HEK293 cells with Mam M/H plasmid are analysed by anti-BU101, anti-Mam, and anti-myc polyclonal and monoclonal antibodies.

FIGURE 7 is a scan of a Western blot from an isoelectric focusing gel (pH 3-10).

FIGURE 8 is a scan of 2 dot blots showing immunorecognition of material by an anti-myc monoclonal antibody. The upper blot shows the fractions from supernatant of the MB8 cells eluting from a Mono Q 5/5 column. The lower blot shows the fractions from supernatant of the Mam M/H transient transfection of HEK293 cells eluting from a Mono Q 5/5 column.

FIGURE 9 is a standard curve for a Superose 12 column showing the relationship between elution volume and molecular weight of protein standards.

FIGURE 10 is a scan of a dot Blot showing immunorecognition of material by an anti-myc monoclonal antibody. The blot shows the fractions from supernatant of the MB8 cells eluting from a Superose 12 column.

FIGURE 11 is a scan of 2 Western blots analysing two tissue extracts and two supernatants with recombinant myc-his tagged Mam and BU101. The upper blot was developed with an anti-BU101 monoclonal antibody and the lower blot was developed with an anti-Mam polyclonal antibody.

FIGURE 12 is a scan of 2 dot blots showing immunorecognition of material by an anti-BU101 polyclonal antibody (upper blot) or an anti-Mam polyclonal antibody (lower blot). Both blots represent the fractions from a breast cancer tissue extract eluting from a Mono Q 5/5 column.

FIGURE 13 is a scan of 2 Western blots showing immunorecognition of material by an anti-BU101 polyclonal antibody (upper blot) or an anti-Mam polyclonal antibody (lower blot). Both blots represent the fractions from a breast cancer tissue extract eluting from a Mono Q 5/5 column.

FIGURE 14 is a scan of 2 dot blots showing immunorecognition of material by an anti-BU101 polyclonal antibody (upper blot) or an anti-Mam polyclonal antibody (lower blot). Both blots represent the fractions from a breast cancer tissue extract eluting from a Superose 12 column.

FIGURE 15 is a scan of a dot blot showing enhanced immunorecognition of myc-his tagged polypeptides using pretreatment protocols.

FIGURE 16 is the BU101 amino acid sequence.

FIGURE 17 is the assembly of BS106 from individual expressed tags.

FIGURE 18A is the BS 106 polynucleotide sequence and 18B is the BS106 polypeptide sequence.

FIGURE 19A, 19B and 19C shows the relative expression of BU101, mammaglobin and BS106, respectively.

FIGURE 20 A-D shows BU101 complexing with mammaglobin.

FIGURE 21 shows correlation between marker expression and clinical and molecular parameters.

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides a new entity, specifically, a multimeric polypeptide complex, wherein at least one copy of BU101 polypeptide (SEQUENCE ID NO 6) and at least one copy of Mammaglobin polypeptide (SEQUENCE ID NO 5) are present, but the complex may contain one or more unknown polypeptides as well. Mammaglobin may be present as a glycoprotein, with sugars attached at asparagine residues located at position 53, and/or position 68, or neither. Furthermore, Mammaglobin may be linked covalently via disulfide bonds to BU101. Both polypeptides contain 3 cysteine residues in their mature form. This disulfide linked heterodimer may constitute one subunit of the complex and it may have interactions with another subunit of identical composition, forming an $\alpha\beta/\alpha\beta$ heterotetramer; or, it may interact with a subunit of nonidentical composition, forming an $\alpha\beta/\alpha'\beta$, or an $\alpha\beta/\alpha\beta'$, or an $\alpha\beta/\alpha'\beta'$ heterotetramer, where α represents BU101, β represents Mammaglobin, α' represents a sequence homologous to but not identical to BU101, and β' represents a sequence homologous to but not identical to Mammaglobin. The BU101 gene may contain a T/C polymorphism at position 254 of the BU101 polynucleotide sequence (SEQUENCE ID NO 2). This polymorphism results in either the amino acid proline (CCG) or the amino acid leucine (CTG) at this position. No biological difference was observed in any experiments described in this invention between the two BU101 nucleotide variants, or in the respectively expressed polypeptides. The multimeric polypeptide complex can be produced by recombinant technology, produced by isolation from natural sources, or produced by synthetic techniques.

The present invention also provides methods for assaying a test sample for this multimeric polypeptide complex which comprises making reagents such as polypeptides, including but not limited to, whole or partial sequences of the component polypeptide

chains of the multimeric polypeptide complex, and antibodies against these antigens. Test samples which may be assayed by the methods provided herein include tissues, cells, body fluids including urine, and secretions.

Portions of the polypeptide sequences are useful as standards or reagents in diagnostic immunoassays, as targets for pharmaceutical screening assays and/or as components or as target sites for various therapies. Monoclonal and polyclonal antibodies directed against at least one epitope contained within these polypeptide sequences are useful as delivery agents for therapeutic agents as well as for diagnostic tests and for screening for diseases or conditions associated with the multimeric polypeptide complex, especially breast cancer.

Techniques for determining amino acid sequence "similarity" are well-known in the art. In general, "similarity" means the exact amino acid to amino acid comparison of two or more polypeptides at the appropriate place, where amino acids are identical or possess similar chemical and/or physical properties such as charge or hydrophobicity. A so-termed "percent similarity" then can be determined between the compared polypeptide sequences. Techniques for determining amino acid sequence identity also are well known in the art and include determining the amino acid sequence and comparing this to a second amino acid sequence. In general, "identity" refers to an exact amino acid to amino acid correspondence of two polypeptide sequences. Two or more amino acid sequences can be compared by determining their "percent identity." The percent identity of two sequences, peptide sequences, is the number of exact matches between two aligned sequences divided by the length of the shorter sequences and multiplied by 100. An approximate alignment for sequences is provided by the local homology algorithm of Smith and Waterman, Advances in Applied Mathematics 2:482-489 (1981). This algorithm can be extended to use with peptide sequences using the scoring matrix developed by Dayhoff, Atlas of Protein Sequences and Structure, M.O. Dayhoff ed., 5 suppl. 3:353-358, National Biomedical Research Foundation, Washington, D.C., USA, and normalized by Gribskov, Nucl. Acids Res. 14(6):6745-6763 (1986). An implementation of this algorithm for peptide sequences is provided by the Genetics Computer Group (Madison, WI) in their BestFit utility application. The default parameters for this method are described in the Wisconsin Sequence Analysis Package Program Manual, Version 8 (1995) (available from Genetics Computer Group, Madison,WI). Other equally suitable programs for

calculating the percent identity or similarity between sequences are generally known in the art.

The compositions and methods described herein will enable the identification of certain markers as indicative of a breast tissue disease or condition; the information obtained therefrom will aid in the detecting, diagnosing, staging, monitoring, prognosticating, in vivo imaging, preventing or treating, or determining diseases or conditions associated with the multimeric polypeptide complex, especially breast cancer. Test methods include, for example, immunoassays which utilize the multimeric polypeptide complex provided herein.

This multimeric polypeptide complex contains unique epitopes which may be found to be immunogenic. These epitopes are believed to be unique to the disease state or condition associated with the multimeric polypeptide complex. It also is thought that the multimeric polypeptide complex is useful as a marker. This marker is either elevated in disease such as breast cancer, altered in disease such as breast cancer, or present as a normal protein complex but appearing in an inappropriate body compartment. The uniqueness of the epitope(s) may be determined by (i) their immunological reactivity and specificity with antibodies directed against proteins and polypeptides encoded by the BU101 gene (SEQUENCE ID NO 2), the Mammaglobin gene (SEQUENCE ID NO 1), the α', or β' genes; and (ii) the absence of cross-reactivity with any other tissue markers. Methods for determining immunological reactivity are well-known and include, but are not limited to, for example, radioimmunoassay (RIA), enzyme-linked immunoabsorbent assay (ELISA), hemagglutination (HA), fluorescence polarization immunoassay (FPIA), chemiluminescent immunoassay (CLIA) and others. Several examples of suitable methods are described herein.

Furthermore, the biological synthesis and assembly of the multimeric polypeptide complex within a cell is highly regulated under normal conditions. Under conditions of disease, the synthesis and assembly of the multimeric polypeptide complex may become deregulated. Deregulation of transcriptional activation may cause an up-regulation or down-regulation of that gene product. Under circumstances where the gene of only one of the component polypeptide chains is up-regulated, increased levels of this gene product may be transcribed and translated into polypeptide. This may cause an accumulation of the single polypeptide independent of other components of the multimeric polypeptide

complex. The measurement of the multimeric polypeptide complex with respect to the free or total amounts of the component polypeptides of the multimeric polypeptide complex is an indication of this deregulation and may indicate a disease such as breast cancer.

Furthermore, deregulation of the synthesis and assembly of the multimeric polypeptide complex may result in overproduction/accumulation of any one of the component polypeptide chains of this complex. Under normal circumstances, the multimeric polypeptide complex is found secreted from the mammalian cell of origin. However, individual component polypeptide chains, independent of other components of the multimeric polypeptide complex, may not undergo the same processing. For example, the expression of BU101 polypeptide, independent of other polypeptides of the multimeric polypeptide complex, may result in a non-secreted form of the polypeptide that is retained inside the cell. The accumulation of free BU101 or other component polypeptide chains inside the cell may result from aberrations in transcription and/or translation, and may be a result of disease such as cancer.

Furthermore, cellular damage and disruption may result in the release of BU101 polypeptide (SEQUENCE ID NO 6) from inside the cell without its multimeric, paired polypeptide chain. This may lead to accumulation of BU101 polypeptide or other component polypeptide chains in the interstitial fluid surrounding the cells. Furthermore, tissue damage and disruption may result in the release of BU101 polypeptide (SEQUENCE ID NO 6) or other component polypeptide chains from the tissue without its respective, paired polypeptide chain. Both cellular and tissue damage/disruption may be a result of disease such as breast cancer.

Unless otherwise stated, the following terms shall have the following meanings:

"Encoded by" refers to a nucleic acid sequence which codes for a polypeptide sequence, wherein the polypeptide sequence or a portion thereof contains an amino acid sequence of at least 3 to 5 amino acids, more preferably at least 8 to 10 amino acids, and even more preferably at least 15 to 20 amino acids from a polypeptide encoded by the nucleic acid sequence. Also encompassed are polypeptide sequences which are immunologically identifiable with a polypeptide encoded by the sequence. Thus, a "polypeptide," "protein," or "amino acid" sequence has at least about 50% identity, preferably about 60% identity, more preferably about 75-85% identity, and most preferably about 90-95% or more identity with an amino acid sequence encoded

by a Mammaglobin gene (SEQUENCE ID NO 1), a BU101 gene (SEQUENCE ID NO 2), an α', or β' gene (i.e., the genes whose encoded polypeptides form the multimeric polypeptide complex). Further, the "polypeptide," "protein," or "amino acid" sequence encoded by the Mammaglobin gene (SEQUENCE ID NO 1), BU101 gene (SEQUENCE ID NO 2), α'gene, or β' gene (i.e., the genes whose encoded polypeptides form the multimeric polypeptide complex) may have at least about 60% similarity, preferably at least about 75% similarity, more preferably about 85% similarity, and most preferably about 95% or more similarity to a polypeptide or amino acid sequence of the Mammaglobin gene (SEQUENCE ID NO 1), BU101 gene (SEQUENCE ID NO 2), α' gene, or β' gene.

A "recombinant polypeptide," "recombinant protein," or "a polypeptide produced by recombinant techniques," which terms may be used interchangeably herein, describes a polypeptide which by virtue of its origin or manipulation is not associated with all or a portion of the polypeptide with which it is associated in nature and/or is linked to a polypeptide other than that to which it is linked in nature. A recombinant or encoded polypeptide or protein is not necessarily translated from a designated nucleic acid sequence. It also may be generated in any manner, including chemical synthesis or expression of a recombinant expression system.
The term "synthetic peptide" as used herein means a polymeric form of amino acids of any length, which may be chemically synthesized by methods well-known to the routineer. These synthetic peptides are useful in various applications.

"Purified polypeptide" or "purified protein" means a polypeptide of interest or fragment thereof which is essentially free of, e.g., contains less than about 50%, preferably less than about 70%, and more preferably less than about 90%, cellular components with which the polypeptide of interest is naturally associated. Methods for purifying polypeptides of interest are known in the art.

The term "isolated" means that the material is removed from its original environment (e.g., the natural environment if it is naturally occurring). For example, a naturally-occurring polynucleotide or polypeptide present in a living animal is not isolated, but the same polynucleotide or DNA or polypeptide, which is separated from some or all of the coexisting materials in the natural system, is isolated. Such polynucleotide could be part of a vector and/or such polynucleotide or polypeptide

could be part of a composition, and still be isolated in that the vector or composition is not part of its natural environment.

"Polypeptide" and "protein" are used interchangeably herein and include all polypeptides as described below. The basic structure of polypeptides is well known and has been described in innumerable textbooks and other publications in the art. In this context, the term is used herein to refer to any peptide or protein comprising two or more amino acids joined to each other in a linear chain by peptide bonds. As used herein, the term refers to both short chains, which also commonly are referred to in the art as peptides, oligopeptides and oligomers, for example, and to longer chains, which generally are referred to in the art as proteins, of which there are many types.

It will be appreciated that polypeptides often contain amino acids other than the 20 amino acids commonly referred to as the 20 naturally occurring amino acids, and that many amino acids, including the terminal amino acids, may be modified in a given polypeptide, either by natural processes, such as processing and other post-translational modifications, but also by chemical modification techniques which are well known to the art. Even the common modifications that occur naturally in polypeptides are too numerous to list exhaustively here, but they are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature, and they are well known to those of skill in the art. Among the known modifications which may be present in polypeptides of the present are, to name an illustrative few, acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid of lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cystine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myrisoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination.

Such modifications are well known to those of skill and have been described in great detail in the scientific literature. Several particularly common modifications, glycosylation, lipid attachment, sulfation, gamma-carboxylation of glutamic acid residues, hydroxylation and ADP-ribosylation, for instance, are described in most

basic texts, such as for instance Proteins – Structure and Molecular Properties, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993). Many detailed reviews are available on this subject, such as, for example, those provided by Wold, F., Posttranslational Protein Modifications: Perspectives and Prospects, pg. 1-12 in Posttranslational Covalent Modification of Proteins, B. C. Johnson, Ed., Academic Press, New York (1983); Seifter et al., Analysis for protein modifications and nonprotein cofactors, Meth. Enzymol. 182: 626-646 (1990) and Rattan et al., Protein synthesis: Posttranslational Modifications and Aging, Ann N. Y. Acad Sci. 663:48-62(1992).

It will be appreciated, as is well known and as noted above, that polypeptides are not always entirely linear. For instance, polypeptides may be branched as a result of ubiquitination, and they may be circular, with or without branching, generally as a result of posttranslational events, including natural processing events and events brought about by human manipulation which do not occur naturally. Circular, branched, and branched circular polypeptides may be synthesized by non-translational natural process and by entirely synthetic methods as well.

Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. In fact, blockage of the amino or carboxyl group in a polypeptide, or both, by a covalent modification, is common in naturally occurring and synthetic polypeptides. For instance, the amino terminal residue of polypeptides made in E. coli, prior to proteolytic processing, almost invariably will be N-formylmethionine.

The modifications that occur in a polypeptide often will be a function of how it is made. For polypeptides made by expressing a cloned gene in a host, for instance, the nature and extent of the modifications in large part will be determined by the host cell posttranslational modification capacity and the modification signals present in the polypeptide amino acid sequence. For instance, as is well known, glycosylation often does not occur in bacterial hosts such as E. coli. Accordingly, when glycosylation is desired, a polypeptide should be expressed in a glycosylating host, generally a eukaryotic cell. Insect cells often carry out the same posttranslational glycosylations as mammalian cells, and, for this reason, insect cell expression systems have been developed to express efficiently mammalian proteins having native patterns of glycosylation. Similar considerations apply to other modifications.

It will be appreciated that the same type of modification may be present in the same or varying degree at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications.

In general, as used herein, the term polypeptide encompasses all such modifications, particularly those that are present in polypeptides synthesized by expressing a polynucleotide in a host cell.

A "multimeric polypeptide complex" and "multimeric polypeptide antigen" and "polypeptide complex" are used interchangeably herein and refer to an entity comprising at least two or more separate individual polypeptide chains. These chains, either identical or different, can be covalently or non-covalently associated. These polypeptide chains can be produced recombinantly, synthesized chemically, or isolated from natural sources. Techniques, known to those in the art, are available for creating a multimeric polypeptide complex from individual chains.

The simplest example of such a multimeric polypeptide complex is a dimer, wherein the individual chains are identical. These chains may be covalently linked, for example, by disulfide bonds, or may be non-covalently associated, for example, by hydrogen bonds and electrostatic interactions. A slightly more complex example is a dimer, wherein the individual chains are non-identical. Again, these chains may be covalently or non-covalently linked. Another level of complexity in a multimeric polypeptide complex would be the increased number of individual chains contributing to the complex, forming trimers, tetramers, pentamers, and higher-order complexes. Again, some chains may interact covalently while others interact non-covalently. Such arrangements may be observed in a crystal structure or solution structure of the protein, and the detailed nature of the interactions that produce the structure become apparent. For example, hemoglobin has 4 individual chains in its complex, 2 of which are of one sequence (α), and 2 of which are of another sequence (β) gene, forming an $\alpha_2\beta_2$ heterotetramer. Another example is E. coli RNA polymerase having the arrangement $\alpha_2\beta\beta'\sigma$, where 2 α chains, 1 β chain, 1 β' chain (homologous to but not identical to β) and 1 σ chain make up the complex. Many more complex cases are known. Some proteins have large numbers of each of several chains, still with a fixed total size and stoichiometry, for example, pyruvate dehydrogenase $[t_{24}(p_2)_{12}(f_2)_{12}]$ which is composed of 24 copies of subunit, t, 12 copies of the homodimer, p_2, and 12 copies of the homodimer, f_2. Others are polymeric structures where the relative

composition may be fixed but the overall size is not, for example, microtubules $[(\alpha\beta)_n]$.

In general, as used herein, the term multimeric polypeptide complex encompasses all such arrangements.

The term "mature" polypeptide refers to a polypeptide which has undergone a complete, post-translational modification appropriate for the subject polypeptide and the cell of origin.

A "fragment" of a specified polypeptide refers to an amino acid sequence which comprises at least about 3-5 amino acids, more preferably at least about 8-10 amino acids, and even more preferably at least about 15-20 amino acids derived from the specified polypeptide.

"Recombinant host cells," "host cells," "cells," "cell lines," "cell cultures," and other such terms denoting microorganisms or higher eukaryotic cell lines cultured as unicellular entities refer to cells which can be, or have been, used as recipients for recombinant vector or other transferred DNA, and include the original progeny of the original cell which has been transfected.

As used herein "replicon" means any genetic element, such as a plasmid, a chromosome or a virus, that behaves as an autonomous unit of polynucleotide replication within a cell.

A "vector" is a replicon in which another polynucleotide segment is attached, such as to bring about the replication and/or expression of the attached segment.

The term "control sequence" refers to a polynucleotide sequence which is necessary to effect the expression of a coding sequence to which it is ligated. The nature of such control sequences differs depending upon the host organism. In prokaryotes, such control sequences generally include a promoter, a ribosomal binding site and terminators; in eukaryotes, such control sequences generally include promoters, terminators and, in some instances, enhancers. The term "control sequence" thus is intended to include at a minimum all components whose presence is necessary for expression, and also may include additional components whose presence is advantageous, for example, leader sequences.

The term "open reading frame" or "ORF" refers to a region of a polynucleotide sequence which encodes a polypeptide. This region may represent a portion of a coding sequence or a total coding sequence.

A "coding sequence" is a polynucleotide sequence which is transcribed into mRNA and translated into a polypeptide when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a translation start codon at the 5' -terminus and a translation stop codon at the 3' -terminus. A coding sequence can include, but is not limited to, mRNA, cDNA and recombinant polynucleotide sequences.

The term "immunologically identifiable with/as" refers to the presence of epitope(s) and polypeptide(s) which also are present in and are unique to the designated polypeptide(s). Immunological identity may be determined by antibody binding and/or competition in binding. These techniques are known to the routineer and also are described herein. The uniqueness of an epitope also can be determined by computer searches of known data banks, such as GenBank, for the polynucleotide sequence which encodes the epitope and by amino acid sequence comparisons with other known proteins.

As used herein, "epitope" means an antigenic determinant of a polypeptide or protein. Conceivably, an epitope can comprise three amino acids in a spatial conformation which is unique to the epitope. Generally, an epitope consists of at least five such amino acids and more usually, it consists of at least eight to ten amino acids. Methods of examining spatial conformation are known in the art and include, for example, x-ray crystallography and two-dimensional nuclear magnetic resonance.

A "conformational epitope" is an epitope that is comprised of a specific juxtaposition of amino acids in an immunologically recognizable structure, such amino acids being present on the same polypeptide in a contiguous or non-contiguous order or present on different polypeptides.

A polypeptide is "immunologically reactive" with an antibody when it binds to an antibody due to antibody recognition of a specific epitope contained within the polypeptide. Immunological reactivity may be determined by antibody binding, more particularly, by the kinetics of antibody binding, and/or by competition in binding using as competitor(s) a known polypeptide(s) containing an epitope against which the antibody is directed. The methods for determining whether a polypeptide is immunologically reactive with an antibody are known in the art.

As used herein, the term "immunogenic polypeptide containing an epitope of interest" means naturally occurring polypeptides of interest or fragments thereof, as

well as polypeptides prepared by other means, for example, by chemical synthesis or the expression of the polypeptide in a recombinant organism.

The term "transfection" refers to the introduction of an exogenous polynucleotide into a prokaryotic or eucaryotic host cell, irrespective of the method used for the introduction. The term "transfection" refers to both stable and transient introduction of the polynucleotide, and encompasses direct uptake of polynucleotides, transformation, transduction, and f-mating. Once introduced into the host cell, the exogenous polynucleotide may be maintained as a non-integrated replicon, for example, a plasmid, or alternatively, may be integrated into the host genome.

"Treatment" refers to prophylaxis and/or therapy.

The term "individual" as used herein refers to vertebrates, particularly members of the mammalian species and includes, but is not limited to, domestic animals, sports animals, primates and humans; more particularly, the term refers to humans.

The term "test sample" refers to a component of an individual's body which is the source of the analyte (such as antibodies of interest or antigens of interest). These components are well known in the art. A test sample is typically anything suspected of containing a target sequence. Test samples can be prepared using methodologies well known in the art such as by obtaining a specimen from an individual and, if necessary, disrupting any cells contained thereby to release target nucleic acids. These test samples include biological samples which can be tested by the methods of the present invention described herein and include human and animal body fluids such as whole blood, serum, plasma, cerebrospinal fluid, sputum, bronchial washing, bronchial aspirates, urine, lymph fluids, and various external secretions of the respiratory, intestinal and genitourinary tracts, tears, saliva, milk, white blood cells, myelomas and the like; biological fluids such as cell culture supernatants; tissue specimens which may be fixed; and cell specimens which may be fixed.

"Purified product" refers to a preparation of the product which has been isolated from the cellular constituents with which the product is normally associated and from other types of cells which may be present in the sample of interest.

"Analyte," as used herein, is the substance to be detected which may be present in the test sample. The analyte can be any substance for which there exists a naturally occurring specific binding member (such as an antibody), or for which a specific binding member can be prepared. Thus, an analyte is a substance that can

bind to one or more specific binding members in an assay. "Analyte" also includes any antigenic substances, haptens, antibodies and combinations thereof. As a member of a specific binding pair, the analyte can be detected by means of naturally occurring specific binding partners (pairs) such as the use of intrinsic factor protein as a member of a specific binding pair for the determination of Vitamin B12, the use of folate-binding protein to determine folic acid, or the use of a lectin as a member of a specific binding pair for the determination of a carbohydrate. The analyte can include a protein, a polypeptide, an amino acid, a nucleotide target and the like. The analyte can be soluble in a body fluid such as blood, blood plasma or serum, urine or the like. The analyte can be in a tissue, either on a cell surface or within a cell. The analyte can be on or in a cell dispersed in a body fluid such as blood, urine, breast aspirate, or obtained as a biopsy sample.

The terms "diseases of the breast," "breast disease," and "condition of the breast" are used interchangeably herein to refer to any disease or condition of the breast including, but not limited to, atypical hyperplasia, fibroadenoma, cystic breast disease, and cancer.

"Breast cancer," as used herein, refers to any malignant disease of the breast including, but not limited to, ductal carcinoma in situ, lobular carcinoma in situ, infiltrating ductal carcinoma, medullary carcinoma, tubular carcinoma, mucinous carcinoma, infiltrating lobular carcinoma, infiltrating comedocarcinoma and inflammatory carcinoma.

An "Expressed Sequence Tag" or "EST" refers to the partial sequence of a cDNA insert which has been made by reverse transcription of mRNA extracted from a tissue followed by insertion into a vector.

A "transcript image" refers to a table or list giving the quantitative distribution of ESTs in a library and represents the genes active in the tissue from which the library was made.

The present invention provides assays which utilize specific binding members. A "specific binding member," as used herein, is a member of a specific binding pair. That is, two different molecules where one of the molecules, through chemical or physical means, specifically binds to the second molecule. Therefore, in addition to antigen and antibody specific binding pairs of common immunoassays, other specific binding pairs can include biotin and avidin, carbohydrates and lectins, complementary nucleotide sequences, effector and receptor molecules, cofactors and enzymes,

enzyme inhibitors, and enzymes and the like. Furthermore, specific binding pairs can include members that are analogs of the original specific binding members, for example, an analyte-analog. Immunoreactive specific binding members include antigens, antigen fragments, antibodies and antibody fragments, both monoclonal and polyclonal and complexes thereof, including those formed by recombinant DNA molecules.

Specific binding members include "specific binding molecules." A "specific binding molecule" intends any specific binding member, particularly an immunoreactive specific binding member. As such, the term "specific binding molecule" encompasses antibody molecules (obtained from both polyclonal and monoclonal preparations), as well as, the following: hybrid (chimeric) antibody molecules (see, for example, Winter, et al., Nature 349:293-299 (1991), and U.S. Patent No. 4,816,567); $F(ab')_2$ and F(ab) fragments; Fv molecules (non-covalent heterodimers, see, for example, Inbar, et al., Proc. Natl. Acad. Sci. USA 69:2659-2662 (1972), and Ehrlich, et al., Biochem. 19:4091-4096 (1980)); single chain Fv molecules (sFv) (see, for example, Huston, et al., Proc. Natl. Acad. Sci. USA 85:5879-5883 (1988)); humanized antibody molecules (see, for example, Riechmann, et al., Nature 332:323-327 (1988), Verhoeyan, et al., Science 239:1534-1536 (1988), and UK Patent Publication No. GB 2,276,169, published 21 September 1994); and, any functional fragments obtained from such molecules, wherein such fragments retain immunological binding properties of the parent antibody molecule.

The term "hapten," as used herein, refers to a partial antigen or non-protein binding member which is capable of binding to an antibody, but which is not capable of eliciting antibody formation unless coupled to a carrier protein.

A "capture reagent," as used herein, refers to an unlabeled specific binding member which is specific either for the analyte as in a sandwich assay, for the indicator reagent or analyte as in a competitive assay, or for an ancillary specific binding member, which itself is specific for the analyte, as in an indirect assay. The capture reagent can be directly or indirectly bound to a solid phase material before the performance of the assay or during the performance of the assay, thereby enabling the separation of immobilized complexes from the test sample.

The "indicator reagent" comprises a "signal-generating compound" ("label") which is capable of generating and generates a measurable signal detectable by external means, conjugated ("attached") to a specific binding member. In addition to

being an antibody member of a specific binding pair, the indicator reagent also can be a member of any specific binding pair, including either hapten-anti-hapten systems such as biotin or anti-biotin, avidin or biotin, a carbohydrate or a lectin, a complementary nucleotide sequence, an effector or a receptor molecule, an enzyme cofactor and an enzyme, an enzyme inhibitor or an enzyme and the like. An immunoreactive specific binding member can be an antibody, an antigen, or an antibody/antigen complex that is capable of binding either to the polypeptide of interest as in a sandwich assay, to the capture reagent as in a competitive assay, or to the ancillary specific binding member as in an indirect assay. When describing probes and probe assays, the term "reporter molecule" may be used. A reporter molecule comprises a signal generating compound as described hereinabove conjugated to a specific binding member of a specific binding pair, such as carbazole or adamantane.

The various "signal-generating compounds" (labels) contemplated include chromagens, catalysts such as enzymes, luminescent compounds such as fluorescein and rhodamine, chemiluminescent compounds such as dioxetanes, acridiniums, phenanthridiniums and luminol, radioactive elements and direct visual labels. Examples of enzymes include alkaline phosphatase, horseradish peroxidase, beta-galactosidase and the like. The selection of a particular label is not critical, but it must be capable of producing a signal either by itself or in conjunction with one or more additional substances.

"Solid phases" ("solid supports") are known to those in the art and include the walls of wells of a reaction tray, test tubes, polystyrene beads, magnetic or non-magnetic beads, nitrocellulose strips, membranes, microparticles such as latex particles, sheep (or other animal) red blood cells and Duracytes® (red blood cells "fixed" by pyruvic aldehyde and formaldehyde, available from Abbott Laboratories, Abbott Park, IL) and others. The "solid phase" is not critical and can be selected by one skilled in the art. Thus, latex particles, microparticles, magnetic or non-magnetic beads, membranes, plastic tubes, walls of microtiter wells, glass or silicon chips, sheep (or other suitable animal's) red blood cells and Duracytes® are all suitable examples. Suitable methods for immobilizing peptides on solid phases include ionic, hydrophobic, covalent interactions and the like. A "solid phase," as used herein, refers to any material which is insoluble, or can be made insoluble by a subsequent reaction. The solid phase can be chosen for its intrinsic ability to attract and

immobilize the capture reagent. Alternatively, the solid phase can retain an additional receptor which has the ability to attract and immobilize the capture reagent. The additional receptor can include a charged substance that is oppositely charged with respect to the capture reagent itself or to a charged substance conjugated to the capture reagent. As yet another alternative, the receptor molecule can be any specific binding member which is immobilized upon (attached to) the solid phase and which has the ability to immobilize the capture reagent through a specific binding reaction. The receptor molecule enables the indirect binding of the capture reagent to a solid phase material before the performance of the assay or during the performance of the assay. The solid phase thus can be a plastic, derivatized plastic, magnetic or non-magnetic metal, glass or silicon surface of a test tube, microtiter well, sheet, bead, microparticle, chip, sheep (or other suitable animal's) red blood cells, Duracytes® and other configurations known to those of ordinary skill in the art.

It is contemplated and within the scope of the present invention that the solid phase also can comprise any suitable porous material with sufficient porosity to allow access by detection antibodies and a suitable surface affinity to bind antigens. Microporous structures generally are preferred, but materials with a gel structure in the hydrated state may be used as well. Such useful solid supports include, but are not limited to, nitrocellulose and nylon. It is contemplated that such porous solid supports described herein preferably are in the form of sheets of thickness from about 0.01 to 0.5 mm, preferably about 0.1 mm. The pore size may vary within wide limits and preferably is from about 0.025 to 15 microns, especially from about 0.15 to 15 microns. The surface of such supports may be activated by chemical processes which cause covalent linkage of the antigen or antibody to the support. The irreversible binding of the antigen or antibody is obtained, however, in general, by adsorption on the porous material by poorly understood hydrophobic forces. Other suitable solid supports are known in the art.

Reagents.

The present invention provides reagents such as a multimeric polypeptide complex or antigen comprising at least one copy of the BU101 polypeptide (SEQUENCE ID NO 6) sequence, at least one copy of the Mammaglobin polypeptide (SEQUENCE ID NO 5), but may contain one or more unknown polypeptides as well, and antibodies specific for this multimeric polypeptide complex. The polypeptides, or antibodies of the present invention may be used to provide information leading to the

detecting, diagnosing, staging, monitoring, prognosticating, in vivo imaging, preventing or treating of, or determining the predisposition to, diseases and conditions of the breast, such as breast cancer.

The present invention relates to a multimeric polypeptide complex which has components with the deduced amino acid sequences as provided in previous applications, as well as one or more unknown polypeptide sequences, as well as fragments, analogs and derivatives of such a multimeric polypeptide complex. The multimeric polypeptide complex of the present invention may be produced recombinantly, purified from natural sources or synthesized. The fragment, derivative or analog of the multimeric polypeptide complex may be one in which one or more of the amino acid residues of any of the component polypeptide chains is substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue) and such substituted amino acid residue may or may not be one encoded by the genetic code; or it may be one in which one or more of the amino acid residues includes a substituent group; or it may be one in which any or all of the chains of the multimeric polypeptide complex is fused with another compound, such as a compound to increase the half-life of the multimeric polypeptide complex (for example, polyethylene glycol); or it may be one in which the additional amino acids are fused to any or all of the chains of the multimeric polypeptide complex, such as a leader or secretory sequence or a sequence which is employed for purification of the multimeric polypeptide complex or a proprotein sequence. Such fragments, derivatives and analogs are within the scope of the present invention. The multimeric polypeptide complex of the present invention are provided preferably in an isolated form and preferably purified.

Thus, any chain of the multimeric polypeptide complex of the present invention may have an amino acid sequence that is identical to that of the naturally occurring polypeptide or that is different by minor variations due to one or more amino acid substitutions. The variation may be a "conservative change" wherein the substituted amino acid has similar structural or chemical properties, e.g., replacement of leucine with isoleucine or threonine with serine. In contrast, variations may include nonconservative changes, e.g., replacement of a glycine with a tryptophan. Similar minor variations may also include amino acid deletions or insertions, or both. Guidance in determining which and how many amino acid residues may be substituted, inserted or deleted without changing biological or immunological activity

may be found using computer programs well known in the art, for example, DNASTAR software (DNASTAR Inc., Madison WI).

Thus, a multimeric polypeptide complex of the present invention may have a composition of at least one copy of the BU101 sequence (α), at least one copy of the Mammaglobin sequence (β), and may or may not have at least one copy of unknown sequences, α'or β'. These components may be present in any ratio.

This invention also provides teachings as to the production of the polypeptides provided herein.

Drug Screening.

The present invention provides a method of screening a plurality of compounds for specific binding to the multimeric polypeptide complex, or any fragment thereof, to identify at least one compound which specifically binds the multimeric polypeptide complex. Such a method comprises the steps of providing at least one compound; combining the multimeric polypeptide complex with each compound under suitable conditions for a time sufficient to allow binding; and detecting the multimeric polypeptide complex binding to each compound.

The polypeptide complex, polypeptides, or peptide fragment(s) employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface or located intracellularly. One method of screening utilizes eukaryotic or prokaryotic host cells which are stably transfected with recombinant nucleic acids which can express the

polypeptide complex, polypeptide or peptide fragment. A drug, compound, or any other agent may be screened against such transfected cells in competitive binding assays. For example, the formation of complexes between a polypeptide and the agent being tested can be measured in either viable or fixed cells.

The present invention thus provides methods of screening for drugs, compounds, or any other agent which can be used to treat diseases associated with the multimeric polypeptide complex. These methods comprise contacting the agent with a polypeptide complex, polypeptide or fragment thereof and assaying for either the presence of a complex between the agent and the polypeptide, or for the presence of a complex between the polypeptide and the cell. In competitive binding assays, the polypeptide typically is labeled. After suitable incubation, free (or uncomplexed) polypeptide or fragment thereof is separated from that present in bound form, and the amount of free or uncomplexed label is used as a measure of the ability of the particular agent to bind to the polypeptide or to interfere with the polypeptide/cell complex.

The present invention also encompasses the use of competitive screening assays in which neutralizing antibodies capable of binding polypeptide specifically compete with a test agent for binding to the polypeptide complex, polypeptide or fragment thereof. In this manner, the antibodies can be used to detect the presence of any polypeptide in the test sample which shares one or more antigenic determinants with a multimeric polypeptide complex as provided herein.

Another technique for screening provides high throughput screening for compounds having suitable binding affinity to at least one polypeptide of the multimeric polypeptide complex disclosed herein. Briefly, large numbers of different small peptide test compounds are synthesized on a solid phase, such as plastic pins or some other surface. The peptide test compounds are reacted with polypeptide and washed. Polypeptide thus bound to the solid phase is detected by methods well-known in the art. Purified polypeptide can also be coated directly onto plates for use in the screening techniques described herein. In addition, non-neutralizing antibodies can be used to capture the polypeptide and immobilize it on the solid support. See, for example, EP 84/03564, published on September 13, 1984, which is incorporated herein by reference.

The goal of rational drug design is to produce structural analogs of biologically active polypeptides of interest or of the small molecules including

agonists, antagonists, or inhibitors with which they interact. Such structural analogs can be used to design drugs which are more active or stable forms of the polypeptide or which enhance or interfere with the function of a polypeptide in vivo. J. Hodgson, Bio/Technology 9:19-21 (1991), incorporated herein by reference.

For example, in one approach, the three-dimensional structure of a polypeptide, or of a polypeptide-inhibitor complex, is determined by x-ray crystallography, by computer modeling or, most typically, by a combination of the two approaches. Both the shape and charges of the polypeptide must be ascertained to elucidate the structure and to determine active site(s) of the molecule. Less often, useful information regarding the structure of a polypeptide may be gained by modeling based on the structure of homologous proteins. In both cases, relevant structural information is used to design analogous polypeptide-like molecules or to identify efficient inhibitors.

Useful examples of rational drug design may include molecules which have improved activity or stability as shown by S. Braxton et al., Biochemistry 31:7796-7801 (1992), or which act as inhibitors, agonists, or antagonists of native peptides as shown by S.B.P. Athauda et al., J Biochem. (Tokyo) 113 (6):742-746 (1993), incorporated herein by reference.

It also is possible to isolate a target-specific antibody selected by an assay as described hereinabove, and then to determine its crystal structure. In principle this approach yields a pharmacophore upon which subsequent drug design can be based. It further is possible to bypass protein crystallography altogether by generating anti-idiotypic antibodies ("anti-ids") to a functional, pharmacologically active antibody. As a mirror image of a mirror image, the binding site of the anti-id is an analog of the original receptor. The anti-id then can be used to identify and isolate peptides from banks of chemically or biologically produced peptides. The isolated peptides then can act as the pharmacophore (that is, a prototype pharmaceutical drug).

A sufficient amount of a recombinant polypeptide complex of the present invention may be made available to perform analytical studies such as X-ray crystallography. In addition, knowledge of the polypeptide amino acid sequences which are derivable from the nucleic acid sequences will provide guidance to those employing computer modeling techniques in place of, or in addition to, x-ray crystallography.

Antibodies specific to a multimeric polypeptide complex (e.g., anti-multimeric polypeptide complex antibodies) further may be used to inhibit the biological action of the polypeptide complex by binding to the polypeptide complex. In this manner, the antibodies may be used in therapy, for example, to treat breast tissue diseases including breast cancer and its metastases.

Further, such antibodies can detect the presence or absence of a multimeric polypeptide complex in a test sample and, therefore, are useful as diagnostic markers for the diagnosis of a breast tissue disease or condition especially breast cancer. Such antibodies may also function as a diagnostic marker for breast tissue disease conditions, such as breast cancer.

The present invention also is directed to antagonists and inhibitors of the polypeptides of the present invention. The antagonists and inhibitors are those which inhibit or eliminate the function of the polypeptide complex. Thus, for example, an antagonist may bind to a polypeptide of the present invention and inhibit or eliminate its function. The antagonist, for example, could be an antibody against the polypeptide which eliminates the activity of a multimeric polypeptide complex by binding a polypeptide, or in some cases the antagonist may be an oligonucleotide. Examples of small molecule inhibitors include, but are not limited to, small peptides or peptide-like molecules.

The antagonists and inhibitors may be employed as a composition with a pharmaceutically acceptable carrier including, but not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol and combinations thereof. Administration of multimeric polypeptide complex inhibitors is preferably systemic. The present invention also provides an antibody which inhibits the action of such a polypeptide complex.

Recombinant Technology.

The present invention provides host cells and expression vectors for the co-expression of BU101 polypeptide (SEQUENCE ID NO 6), Mammaglobin polypeptide (SEQUENCE ID NO 5), and unknown α' and/or β' polypeptides, and methods for the production of the multimeric polypeptide complex they encode. Such methods comprise culturing the host cells under conditions suitable for the expression of BU101 polypeptide (SEQUENCE ID NO 6), Mammaglobin polypeptide (SEQUENCE ID NO 5), and unknown α' and/or β' polypeptides, and recovering the

multimeric polypeptide(s) and/or multimeric polypeptide complex from the cell culture.

The present invention also provides vectors which individually encode BU101, Mammaglobin, and unknown α' and/or β' polypeptide(s) of the present invention; host cells which are genetically engineered with vectors of the present invention; and, the production of all, or any, of the subunits of the multimeric polypeptide complex of the present invention by recombinant techniques.

Host cells are genetically engineered (transfected, transduced or transformed) with vectors which may be cloning vectors or expression vectors. The vector may be in the form of a plasmid, a viral particle, a phage, etc. The engineered host cells can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transfected cells, or amplifying BU101 (SEQUENCE ID NO 2), Mammaglobin (SEQUENCE ID NO 1), or unknown α' and/or β' gene(s). The culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

The polynucleotides may be employed for producing a polypeptide or polypeptide complex(es) by recombinant techniques. Thus, the polynucleotide sequence may be included in any one of a variety of expression vehicles, in particular, vectors or plasmids for expressing a polypeptide. Such vectors include chromosomal, nonchromosomal and synthetic DNA sequences, e.g., derivatives of SV40; bacterial plasmids; phage DNA; yeast plasmids; vectors derived from combinations of plasmids and phage DNA, viral DNA such as vaccinia, adenovirus, fowl pox virus and pseudorabies. However, any other plasmid or vector may be used so long as it is replicable and viable in the host.

The appropriate DNA sequence may be inserted into the vector by a variety of procedures. In general, the DNA sequence is inserted into appropriate restriction endonuclease sites by procedures known in the art. Such procedures and others are deemed to be within the scope of those skilled in the art. The DNA sequence in the expression vector is operatively linked to an appropriate expression control sequence(s) (promoter) to direct mRNA synthesis. Representative examples of such promoters include, but are not limited to, the LTR or the SV40 promoter, the E. coli lac or trp, the phage lambda P sub L promoter and other promoters known to control expression of genes in prokaryotic or eukaryotic cells or their viruses. The expression

vector also contains a ribosome binding site for translation initiation and a transcription terminator. The vector may also include appropriate sequences for amplifying expression. In addition, the expression vectors preferably contain a gene to provide a phenotypic trait for selection of transfected host cells such as dihydrofolate reductase or neomycin resistance for eukaryotic cell culture, or such as tetracycline or ampicillin resistance in E. coli.

The vector containing the appropriate DNA sequence as hereinabove described, as well as an appropriate promoter or control sequence, may be employed to transfect an appropriate host to permit the host to express the protein. As representative examples of appropriate hosts, there may be mentioned: bacterial cells, such as E. coli, Salmonella typhimurium; Streptomyces sp.; fungal cells, such as yeast; insect cells, such as Drosophila and Sf9; animal cells, such as CHO, COS or Bowes melanoma; plant cells, etc. The selection of an appropriate host is deemed to be within the scope of those skilled in the art from the teachings provided herein.

More particularly, the present invention also includes recombinant constructs comprising one or more of the sequences as broadly described above. The constructs comprise a vector, such as a plasmid or viral vector, into which a sequence of the invention has been inserted, in a forward or reverse orientation. In a preferred aspect of this embodiment, the construct further comprises regulatory sequences including, for example, a promoter, operably linked to the sequence. Large numbers of suitable vectors and promoters are known to those of skill in the art and are commercially available. The following vectors are provided by way of example. Bacterial: pINCY (Incyte Pharmaceuticals Inc., Palo Alto, CA), pSPORT1 (Life Technologies, Gaithersburg, MD), pQE70, pQE60, pQE-9 (Qiagen) pBs, phagescript, psiX174, pBluescript SK, pBsKS, pNH8a, pNH16a, pNH18a, pNH46a (Stratagene); pTrc99A, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia); Eukaryotic: pWLneo, pSV2cat, pOG44, pXT1, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia). However, any other plasmid or vector may be used as long as it is replicable and viable in the host.

Plasmid pINCY is generally identical to the plasmid pSPORT1 (available from Life Technologies, Gaithersburg, MD) with the exception that it has two modifications in the polylinker (multiple cloning site). These modifications are (1) it lacks a HindIII restriction site and (2) its EcoRI restriction site lies at a different location. pINCY is created from pSPORT1 by cleaving pSPORT1 with both HindIII

and EcoRI and replacing the excised fragment of the polylinker with synthetic DNA fragments (SEQUENCE ID NOS 3-4). This replacement may be made in any manner known to those of ordinary skill in the art. For example, the two nucleotide sequences, SEQUENCE ID NOS 3-4, may be generated synthetically with 5' terminal phosphates, mixed together, and then ligated under standard conditions for performing staggered end ligations into the pSPORT1 plasmid cut with HindIII and EcoRI. Suitable host cells (such as E. coli DH5α cells) then are transfected with the ligated DNA and recombinant clones are selected for ampicillin resistance. Plasmid DNA then is prepared from individual clones and subjected to restriction enzyme analysis or DNA sequencing in order to confirm the presence of insert sequences in the proper orientation. Other cloning strategies known to the ordinary artisan also may be employed.

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, SP6, T7, gpt, lambda P sub R, P sub L and trp. Eukaryotic promoters include cytomegalovirus (CMV) immediate early, herpes simplex virus (HSV) thymidine kinase, early and late SV40, LTRs from retroviruses and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art.

In a further embodiment, the present invention provides host cells containing the above-described construct. The host cell can be a higher eukaryotic cell, such as a mammalian cell, or a lower eukaryotic cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-Dextran mediated transfection, or electroporation [L. Davis et al., Basic Methods in Molecular Biology, 2nd edition, Appleton and Lang, Paramount Publishing, East Norwalk, CT (1994)].

The constructs in host cells can be used in a conventional manner to produce the gene product encoded by the recombinant sequence. Alternatively, the polypeptides of the invention can be synthetically produced by conventional peptide synthesizers.

Recombinant proteins can be expressed in mammalian cells, yeast, bacteria, or other cells, under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA

constructs. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook et al., Molecular Cloning: A Laboratory Manual, Second Edition, (Cold Spring Harbor, NY, 1989), which is hereby incorporated by reference.

Transcription of a DNA encoding the polypeptide(s) of the present invention by higher eukaryotes is increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp, that act on a promoter to increase its transcription. Examples include the SV40 enhancer on the late side of the replication origin (bp 100 to 270), a cytomegalovirus early promoter enhancer, a polyoma enhancer on the late side of the replication origin and adenovirus enhancers.

Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transfection of the host cell, e.g., the ampicillin resistance gene of E. coli and S. cerevisiae TRP1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), alpha factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an N-terminal identification peptide imparting desired characteristics, e.g., stabilization or simplified purification of expressed recombinant product.

Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transfection include E. coli, Bacillus subtilis, Salmonella typhimurium and various species within the genera Pseudomonas, Streptomyces and Staphylococcus, although others may also be employed as a routine matter of choice.

Useful expression vectors for bacterial use comprise a selectable marker and bacterial origin of replication derived from plasmids comprising genetic elements of the well-known cloning vector pBR322 (ATCC 37017). Other vectors include but are not limited to PKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM1 (Promega Biotec, Madison, WI). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed.

Following transfection of a suitable host and growth of the host to an appropriate cell density, the selected promoter is derepressed by appropriate means (e.g., temperature shift or chemical induction), and cells are cultured for an additional period. Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification. Microbial cells employed in expression of proteins can be disrupted by any convenient method including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents. Such methods are well-known to the ordinary artisan.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts described by Gluzman, Cell 23:175 (1981), and other cell lines capable of expressing a compatible vector, such as the C127, HEK-293, 3T3, CHO, HeLa and BHK cell lines. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and enhancer and also any necessary ribosome binding sites, polyadenylation sites, splice donor and acceptor sites, transcriptional termination sequences and 5' flanking nontranscribed sequences. DNA sequences derived from the SV40 viral genome, for example, SV40 origin, early promoter, enhancer, splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements. Representative, useful vectors include pRc/CMV and pcDNA3 (available from Invitrogen, San Diego, CA).

Polypeptides are recovered and purified from recombinant cell cultures by known methods including affinity chromatography, ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, hydroxyapatite chromatography or lectin chromatography. It is preferred to have low concentrations (approximately 0.1-5 mM) of calcium ion present during purification [Price, et al., J. Biol. Chem. 244:917 (1969)]. Protein refolding steps can be used, as

necessary, in completing configuration of the polypeptide. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps.

Thus, polypeptides of the present invention may be naturally purified products expressed from a high expressing cell line, or a product of chemical synthetic procedures, or produced by recombinant techniques from a prokaryotic or eukaryotic host (for example, by bacterial, yeast, higher plant, insect and mammalian cells in culture). Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated with mammalian or other eukaryotic carbohydrates or may be non-glycosylated. The polypeptides of the invention may also include a methionine residue as initial amino acid.

Plasmids containing cDNAs can be constructed from available plasmids in accord with published, known procedures. In addition, equivalent plasmids to those described are known in the art and will be apparent to one of ordinary skill in the art. The cDNA can be shuttled into other vectors known to be useful for expression of protein in specific hosts. Oligonucleotide primers containing cloning sites and segments of DNA sufficient to hybridize to stretches at both ends of the target cDNA can be synthesized chemically by standard methods. These primers can then be used to amplify the desired gene segments by PCR. The resulting new gene segments can be digested with appropriate restriction enzymes under standard conditions and isolated by gel electrophoresis. Alternately, similar gene segments can be produced by digestion of the cDNA with appropriate restriction enzymes and filling in the missing gene segments with chemically synthesized oligonucleotides. Segments of the coding sequence from more than one gene can be ligated together and cloned in appropriate vectors to optimize expression of recombinant sequence.

Suitable expression hosts for such chimeric molecules include, but are not limited to, mammalian cells, such as Chinese Hamster Ovary (CHO) and human embryonic kidney (HEK) 293 cells, insect cells, such as Sf9 cells, yeast cells, such as Saccharomyces cerevisiae and bacteria, such as E. coli. For each of these cell systems, a useful expression vector may also include an origin of replication to allow propagation in bacteria and a selectable marker such as the beta-lactamase antibiotic resistance gene to allow selection in bacteria. In addition, the vectors may include a second selectable marker, such as the neomycin phosphotransferase gene, to allow selection in transfected eukaryotic host cells. Vectors for use in eukaryotic expression

hosts may require the addition of 3' poly A tail if the sequence of interest lacks poly A.

Additionally, the vector may contain promoters or enhancers which increase gene expression. Such promoters are host specific and include, but are not limited to, MMTV, SV40, or metallothionine promoters for CHO cells; trp, lac, tac or T7 promoters for bacterial hosts; or alpha factor, alcohol oxidase or PGH promoters for yeast. Adenoviral vectors with or without transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to drive protein expression in mammalian cell lines. Once homogeneous cultures of recombinant cells are obtained, large quantities of recombinantly produced protein can be recovered from the conditioned medium and analyzed using chromatographic methods well known in the art. An alternative method for the production of large amounts of secreted protein involves the transfection of mammalian embryos and the recovery of the recombinant protein from milk produced by transgenic cows, goats, sheep, etc. Polypeptides and closely related molecules may be expressed recombinantly in such a way as to facilitate protein purification. One approach involves expression of a chimeric protein which includes one or more additional polypeptide domains not naturally present on human polypeptides. Such purification-facilitating domains include, but are not limited to, metal-chelating peptides such as histidine-tryptophan domains that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp, Seattle, WA). The inclusion of a cleavable linker sequence such as Factor XA or enterokinase from Invitrogen (San Diego, CA) between the polypeptide sequence and the purification domain may be useful for recovering the polypeptide.

Immunoassays.

Polypeptides, multimeric polypeptide complexes, including fragments, derivatives, and analogs thereof, or cells expressing such polypeptides or multimeric polypeptide complexes, can be utilized in a variety of assays, many of which are described herein, for the detection of antibodies to breast tissue. They also can be used as immunogens to produce antibodies. These antibodies can be, for example, polyclonal or monoclonal antibodies, chimeric, single chain and humanized antibodies, as well as Fab fragments, or the product of an Fab expression library.

Various procedures known in the art may be used for the production of such antibodies and fragments.

For example, antibodies generated against a multimeric polypeptide complex can be obtained by direct injection of the multimeric polypeptide complex into an animal or by administering the multimeric polypeptide complex to an animal such as a mouse, rabbit, goat or human. A mouse, rabbit or goat is preferred. The multimeric polypeptide complex is composed of a group of sequences consisting of BU101 (SEQUENCE ID NO 6), Mammaglobin (SEQUENCE ID NO 5), and unknown α' and/or β' polypeptide(s), and fragments thereof. The antibody so obtained then binds the multimeric polypeptide complex itself. In this manner, a sequence encoding only a fragment of the multimeric polypeptide complex or any of its component polypeptides can be used to generate antibodies that bind the native polypeptide complex. Such antibodies then can be used to isolate the multimeric polypeptide complex from test samples such as tissue suspected of containing that multimeric polypeptide complex. For preparation of monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures can be used. Examples include the hybridoma technique as described by Kohler and Milstein, Nature 256:495-497 (1975), the trioma technique, the human B-cell hybridoma technique as described by Kozbor et al., Immun. Today 4:72 (1983) and the EBV-hybridoma technique to produce human monoclonal antibodies as described by Cole et al., in Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc, New York, NY, pp. 77-96 (1985). Techniques described for the production of single chain antibodies can be adapted to produce single chain antibodies to immunogenic polypeptide products of this invention. See, for example, U.S. Patent No. 4,946,778, which is incorporated herein by reference.

The monoclonal antibodies or fragments thereof can be provided individually to detect antigens of the multimeric polypeptide complex. Combinations of the monoclonal antibodies (and fragments thereof) provided herein also may be used together as components in a mixture or "cocktail" wherein at least one antibody which binds to the multimeric polypeptide complex of the invention, along with antibodies which specifically bind to other regions of the multimeric polypeptide complex are present, each antibody having different binding specificities. For example, one monoclonal antibody may recognize a shared epitope wherein that epitope is derived from components of two or more different polypeptides. In this case, the epitope

would be specific for a multimeric complex wherein the polypeptides were present but would not bind to the individual, isolated polypeptides. Another monoclonal antibody may recognize an epitope specified within a single polypeptide sequence. In this case, the epitope may be present in both the individual, isolated polypeptide as well as in the multimeric polypeptide complex. Another monoclonal antibody may recognize an epitope specified within a single polypeptide sequence. In this case, the epitope may be present in the individual, isolated polypeptide but not in the multimeric polypeptide complex. The epitope may be buried or be conformationally distinct as a result of complexation with other polypeptides. Thus, this cocktail can include the monoclonal antibodies of the invention which are directed to any single antigenic determinant of the multimeric polypeptides disclosed herein and other monoclonal antibodies specific to other antigenic determinants of these antigens or other related proteins.

The polyclonal antibody or fragment thereof which can be used in the assay formats should specifically bind to the multimeric polypeptide complex of the present invention or any of the component polypeptides of this complex, or fragments thereof, additionally used in the assay. The polyclonal antibody used preferably is of mammalian origin such as, human, goat, rabbit or sheep polyclonal antibody which binds the multimeric polypeptide complex. Most preferably, the polyclonal antibody is of rabbit origin. The polyclonal antibodies used in the assays can be used either alone or as a cocktail of polyclonal antibodies. Since the cocktails used in the assay formats are comprised of either monoclonal antibodies or polyclonal antibodies having different binding specificity to the multimeric polypeptide complex, they are useful for the detecting, diagnosing, staging, monitoring, prognosticating, in vivo imaging, preventing or treating, or determining the predisposition to, diseases and conditions of the breast, such as breast cancer.

Various assay formats may utilize the antibodies of the present invention, including "sandwich" immunoassays. For example, the antibodies of the present invention, or fragments thereof, can be employed in various assay systems to determine the presence, if any, of the multimeric polypeptide antigen in a test sample. For example, in a first assay format, a polyclonal or monoclonal antibody or fragment thereof, or a combination of these antibodies, which has been coated on a solid phase, is contacted with a test sample, to form a first mixture. This first mixture is incubated for a time and under conditions sufficient to form antigen/antibody complexes. Then,

an indicator reagent comprising a monoclonal or a polyclonal antibody or a fragment thereof, or a combination of these antibodies, to which a signal generating compound has been attached, is contacted with the antigen/antibody complexes to form a second mixture. This second mixture then is incubated for a time and under conditions sufficient to form antibody/antigen/antibody complexes. The presence of antigen in the test sample and captured on the solid phase, if any, is determined by detecting the measurable signal generated by the signal generating compound. The amount of antigen present in the test sample is proportional to the signal generated.

In another example of a sandwich immunoassay, the antibodies of the present invention, or fragments thereof, can be employed in various assay systems to determine the presence, if any, of the individual, isolated polypeptides that constitute the multimeric polypeptide complex. In this case, the antibodies utilized bind to an epitope that is present in the individual, isolated polypeptide but is not available for binding in the multimeric polypeptide complex.

In another example of a sandwich immuoassay, the antibodies of the present invention, or fragments thereof, can be employed in various assay systems to determine the presence, if any, of both the individual, isolated polypeptides that constitute the multimeric polypeptide complex and the multimeric polypeptide complex itself. In this case, the antibodies utilized bind to an epitope that is present in the individual, isolated polypeptide and in the multimeric polypeptide complex.

Measurements of these different antigens, specifically, the multimeric polypeptide complex (bound), the individual, isolated polypeptides (free), and both the multimeric polypeptide complex and the individual, isolated polypeptides (total), and ratios thereof, may be useful for the detecting, diagnosing, staging, monitoring, prognosticating, in vivo imaging, preventing or treating, or determining the predisposition to diseases and conditions of the breast, such as breast cancer. See, for example, International Publication Number WO 92/01936, which is incorporated herein by reference.

In an alternative assay format, a mixture is formed by contacting: (1) a polyclonal antibody, monoclonal antibody, or fragment thereof, which specifically binds to a multimeric polypeptide antigen and/or one of its component polypeptide chains such that measurements of free, bound, or total can be made, or a combination of such antibodies bound to a solid support; (2) the test sample; and (3) an indicator

reagent comprising a monoclonal antibody, polyclonal antibody, or fragment thereof, which specifically binds to a different epitope of the multimeric polypeptide antigen and/or one of its component chains (or a combination of these antibodies) to which a signal generating compound is attached. This mixture is incubated for a time and under conditions sufficient to form antibody/antigen/antibody complexes. The presence, if any, of the multimeric polypeptide antigen and/or one of its component polypeptide chains present in the test sample and captured on the solid phase is determined by detecting the measurable signal generated by the signal generating compound. The amount of antigen present in the test sample is proportional to the signal generated.

In another assay format, antibodies coated on solid phases or labeled with detectable labels are then allowed to compete with those present in a patient sample (if any) for a limited amount of antigen. A reduction in binding of the polyclonal or monoclonal antibodies is an indication of the presence of antigen in the patient sample. The presence of antibodies against the antigen indicates the presence of breast tissue disease, especially breast cancer, in the patient.

In yet another detection method, each of the monoclonal or polyclonal antibodies of the present invention can be employed in the detection of multimeric polypeptide antigens including the multimeric polypeptide complex and/or one of its component polypeptide chains in tissue sections, as well as in cells, by immunohistochemical analysis. The tissue sections can be cut from either frozen or chemically fixed samples of tissue. If the antigens are to be detected in cells, the cells can be isolated from blood, urine, breast aspirates, or other bodily fluids. The cells may be obtained by biopsy, either surgical or by needle. The cells can be isolated by centrifugation or magnetic attraction after labeling with magnetic particles or ferrofluids so as to enrich a particular fraction of cells for staining with the antibodies of the present invention. Cytochemical analysis wherein these antibodies are labeled directly (with, for example, fluorescein, colloidal gold, horseradish peroxidase, alkaline phosphatase, etc.) or are labeled by using secondary labeled anti-species antibodies (with various labels as exemplified herein) to track the histopathology of disease also are within the scope of the present invention.

In addition, these monoclonal antibodies can be bound to matrices similar to CNBr-activated Sepharose and used for the affinity purification of the multimeric

polypeptide complex and/or one of its component chains from cell cultures or biological tissues such as to purify recombinant and native protein.

The monoclonal antibodies of the invention also can be used for the generation of chimeric antibodies for therapeutic use, or other similar applications.

It is contemplated and within the scope of the present invention that the multimeric polypeptide antigen may be detectable in assays by use of a recombinant antigen as well as by use of a synthetic peptide or purified peptide, which peptide comprises an amino acid sequence of any component polypeptide chain of the multimeric polypeptide complex. The amino acid sequence of such a polypeptide is selected from the group consisting of BU101 polypeptide (SEQUENCE ID NO 6), Mammaglobin polypeptide (SEQUENCE ID NO 5), an unknown α' and/or β' polypeptide sequence, and fragments thereof. It also is within the scope of the present invention that different synthetic, recombinant or purified peptides, identifying different epitopes of the multimeric polypeptide complex, can be used in combination in an assay for the detecting, diagnosing, staging, monitoring, prognosticating, in vivo imaging, preventing or treating, or determining the predisposition to diseases and conditions of the breast, such as breast cancer. In this case, all of these peptides or polypeptides can be coated onto one solid phase; or each separate peptide or polypeptide may be coated onto separate solid phases, such as microparticles, and then combined to form a mixture of peptides or polypeptides which can be later used in assays. Furthermore, it is contemplated that multiple peptides or polypeptides which define epitopes from different antigens may be used for the detection, diagnosis, staging, monitoring, prognosis, prevention or treatment of, or determining the predisposition to, diseases and conditions of the breast, such as breast cancer. Peptides or polypeptides coated on solid phases or labeled with detectable labels are then allowed to compete with those present in a patient sample (if any) for a limited amount of antibody. A reduction in binding of the synthetic, recombinant, or purified peptides to the antibody (or antibodies) is an indication of the presence of the multimeric polypeptide antigen in the patient sample. The presence of the multimeric polypeptide antigen indicates the presence of breast tissue disease, especially breast cancer, in the patient. Variations of assay formats are known to those of ordinary skill in the art and many are discussed herein below.

In another assay format, one or a combination of at least two polypeptides, peptides, or the multimeric polypeptide complex of the invention can be employed as

a competitive probe for the detection of the multimeric polypeptide antigen. For example, antibodies to the multimeric polypeptide complex such as the monoclonal and polyclonal antibodies disclosed herein, either alone or in combination, are coated on a solid phase. A test sample suspected of containing the multimeric polypeptide antigen then is incubated with an indicator reagent comprising a signal generating compound and at least one monoclonal antibody of the invention for a time and under conditions sufficient to form antigen/antibody complexes of either the test sample and indicator reagent bound to the solid phase or the indicator reagent bound to the solid phase. The reduction in binding of the monoclonal antibody to the solid phase can be quantitatively measured.

In another assay format, the presence of anti-multimeric polypeptide antibody and/or multimeric polypeptide antigen can be detected in a simultaneous assay, as follows. A test sample is simultaneously contacted with a capture reagent of a first analyte, wherein said capture reagent comprises a first binding member specific for a first analyte attached to a solid phase and a capture reagent for a second analyte, wherein said capture reagent comprises a first binding member for a second analyte attached to a second solid phase, to thereby form a mixture. This mixture is incubated for a time and under conditions sufficient to form capture reagent/first analyte and capture reagent/second analyte complexes. These so-formed complexes then are contacted with an indicator reagent comprising a member of a binding pair specific for the first analyte labeled with a signal generating compound and an indicator reagent comprising a member of a binding pair specific for the second analyte labeled with a signal generating compound to form a second mixture. This second mixture is incubated for a time and under conditions sufficient to form capture reagent/first analyte/indicator reagent complexes and capture reagent/second analyte/indicator reagent complexes. The presence of one or more analytes is determined by detecting a signal generated in connection with the complexes formed on either or both solid phases as an indication of the presence of one or more analytes in the test sample. In this assay format, recombinant antigens derived from the expression systems disclosed herein may be utilized, as well as monoclonal antibodies produced from the proteins derived from the expression systems as disclosed herein. For example, in this assay system, the multimeric polypeptide antigen can be the first analyte. Such assay systems are described in greater detail in EP Publication No. 0473065.

In yet other assay formats, the polypeptides disclosed herein may be utilized to detect the presence of antibody against the multimeric polypeptide antigen in test samples. For example, a test sample is incubated with a solid phase to which at least one polypeptide such as a recombinant protein or synthetic peptide has been attached. The polypeptide is selected from the group consisting of BU101 (SEQUENCE ID NO 6), Mammaglobin (SEQUENCE ID NO 5), α' polypeptide, β'' polypeptide, and fragments thereof. These are reacted for a time and under conditions sufficient to form antigen/antibody complexes. Following incubation, the antigen/antibody complex is detected. Indicator reagents may be used to facilitate detection, depending upon the assay system chosen.

In another assay format, a test sample is contacted with a solid phase to which a recombinant protein produced as described herein is attached, and also is contacted with a monoclonal or polyclonal antibody specific for the protein, which preferably has been labeled with an indicator reagent. After incubation for a time and under conditions sufficient for antibody/antigen complexes to form, the solid phase is separated from the free phase, and the label is detected in either the solid or free phase as an indication of the presence of antibody against the multimeric polypeptide antigen.

Other assay formats utilizing the recombinant antigens disclosed herein are contemplated. These include contacting a test sample with a solid phase to which at least one antigen from a first source has been attached, incubating the solid phase and test sample for a time and under conditions sufficient to form antigen/antibody complexes, and then contacting the solid phase with a labeled antigen, which antigen is derived from a second source different from the first source. For example, a recombinant protein derived from a first source such as E. coli is used as a capture antigen on a solid phase, a test sample is added to the so-prepared solid phase, and following standard incubation and washing steps as deemed or required, a recombinant protein derived from a different source (i.e., non-E. coli) is utilized as a part of an indicator reagent which subsequently is detected. Likewise, combinations of a recombinant antigen on a solid phase and synthetic peptide in the indicator phase also are possible. Any assay format which utilizes an antigen specific for the multimeric polypeptide complex produced or derived from a first source as the capture antigen and an antigen specific for the multimeric polypeptide complex from a different second source is contemplated. Thus, various combinations of

recombinant antigens, as well as the use of synthetic peptides, purified proteins and the like, are within the scope of this invention. Assays such as this and others are described in U.S. Patent No. 5,254,458, which enjoys common ownership and is incorporated herein by reference.

Other embodiments which utilize various other solid phases also are contemplated and are within the scope of this invention. For example, ion capture procedures for immobilizing an immobilizable reaction complex with a negatively charged polymer (described in EP publication No. 0 326 100 and EP publication No. 0 406 473), can be employed according to the present invention to effect a fast solution-phase immunochemical reaction. An immobilizable immune complex is separated from the rest of the reaction mixture by ionic interactions between the negatively charged poly-anion/immune complex and the previously treated, positively charged porous matrix and detected by using various signal generating systems previously described, including those described in chemiluminescent signal measurements as described in EPO Publication No. 0 273,115.

Also, the methods of the present invention can be adapted for use in systems which utilize microparticle technology including automated and semi-automated systems wherein the solid phase comprises a microparticle (magnetic or non-magnetic). Such systems include those described in, for example, published EPO applications Nos. EP 0 425 633 and EP 0 424 634, respectively.

The use of scanning probe microscopy (SPM) for immunoassays also is a technology to which the monoclonal antibodies of the present invention are easily adaptable. In scanning probe microscopy, particularly in atomic force microscopy, the capture phase, for example, at least one of the monoclonal antibodies of the invention, is adhered to a solid phase and a scanning probe microscope is utilized to detect antigen/antibody complexes which may be present on the surface of the solid phase. The use of scanning tunneling microscopy eliminates the need for labels which normally must be utilized in many immunoassay systems to detect antigen/antibody complexes. The use of SPM to monitor specific binding reactions can occur in many ways. In one embodiment, one member of a specific binding partner (analyte specific substance which is the monoclonal antibody of the invention) is attached to a surface suitable for scanning. The attachment of the analyte specific substance may be by adsorption to a test piece which comprises a solid phase of a plastic or metal surface, following methods known to those of ordinary skill in the art.

Or, covalent attachment of a specific binding partner (analyte specific substance) to a test piece which test piece comprises a solid phase of derivatized plastic, metal, silicon, or glass may be utilized. Covalent attachment methods are known to those skilled in the art and include a variety of means to irreversibly link specific binding partners to the test piece. If the test piece is silicon or glass, the surface must be activated prior to attaching the specific binding partner. Also, polyelectrolyte interactions may be used to immobilize a specific binding partner on a surface of a test piece by using techniques and chemistries. The preferred method of attachment is by covalent means. Following attachment of a specific binding member, the surface may be further treated with materials such as serum, proteins, or other blocking agents to minimize non-specific binding. The surface also may be scanned either at the site of manufacture or point of use to verify its suitability for assay purposes. The scanning process is not anticipated to alter the specific binding properties of the test piece.

While the present invention discloses the preference for the use of solid phases, it is contemplated that the reagents such as antibodies, proteins and peptides of the present invention can be utilized in non-solid phase assay systems. These assay systems are known to those skilled in the art, and are considered to be within the scope of the present invention.

It is contemplated that the reagent employed for the assay can be provided in the form of a test kit with one or more containers such as vials or bottles, with each container containing a separate reagent such as a probe, primer, monoclonal antibody or a cocktail of monoclonal antibodies, or a polypeptide (e.g. recombinantly, synthetically produced or purified) employed in the assay. The polypeptide is selected from the group consisting of BU101 (SEQUENCE ID NO 6), Mammaglobin (SEQUENCE ID NO 5), α' polypeptide, or β' polypeptide, and fragments thereof. Other components such as buffers, controls and the like, known to those of ordinary skill in art, may be included in such test kits. It also is contemplated to provide test kits which have means for collecting test samples comprising accessible body fluids or waste products, e.g., blood, urine, saliva and stool. Such tools useful for collection ("collection materials") include lancets and absorbent paper or cloth for collecting and stabilizing blood; swabs for collecting and stabilizing saliva; cups for collecting and stabilizing urine or stool samples. Collection materials, papers, cloths, swabs, cups and the like, may optionally be treated to avoid denaturation or irreversible adsorption

of the sample. The collection materials also may be treated with or contain preservatives, stabilizers or antimicrobial agents to help maintain the integrity of the specimens. Test kits designed for the collection, stabilization and preservation of test specimens obtained by surgery or needle biopsy are also useful. It is contemplated that all kits may be configured in two components which can be provided separately; one component for collection and transport of the specimen and the other component for the analysis of the specimen. The collection component, for example, can be provided to the open market user while the components for analysis can be provided to others such as laboratory personnel for determination of the presence, absence or amount of analyte. Further, kits for the collection, stabilization and preservation of test specimens may be configured for use by untrained personnel and may be available in the open market for use at home with subsequent transportation to a laboratory for analysis of the test sample.

In Vivo Antibody Use.

Antibodies of the present invention can be used in vivo; that is, they can be injected into patients suspected of having or having diseases of the breast for diagnostic or therapeutic uses. The use of antibodies for in vivo diagnosis is well known in the art. Sumerdon et al., Nucl. Med. Biol 17:247-254 (1990) have described an optimized antibody-chelator for the radioimmunoscintographic imaging of carcinoembryonic antigen (CEA) expressing tumors using Indium-111 as the label. Griffin et al., J Clin Onc 9:631-640 (1991) have described the use of this agent in detecting tumors in patients suspected of having recurrent colorectal cancer. The use of similar agents with paramagnetic ions as labels for magnetic resonance imaging is know in the art (R. B. Lauffer, Magnetic Resonance in Medicine 22:339-342 (1991). It is anticipated that antibodies directed against the multimeric polypeptide antigen can be injected into patients suspected of having a disease of the breast such as breast cancer for the purpose of diagnosing or staging the disease status of the patient. The label used will depend on the imaging modality chosen. Radioactive labels such as Indium-111, Technetium-99m, or Iodine-131 can be used for planar scans or single photon emission computed tomography (SPECT). Positron emitting labels such as Fluorine-19 can also be used for positron emission tomography (PET). For MRI, paramagnetic ions such as Gadolinium (III) or Manganese (II) can be used. Localization of the label within the breast or external to the breast may allow

determination of spread of the disease. The amount of label within the breast may allow determination of the presence or absence of cancer of the breast.

For patients known to have a disease of the breast, injection of an antibody directed against the multimeric polypeptide antigen may have therapeutic benefit. The antibody may exert its effect without the use of attached agents by binding to the multimeric polypeptide antigen expressed on or in the tissue or organ. Alternatively, the antibody may be conjugated to cytotoxic agents such as drugs, toxins, or radionuclides to enhance its therapeutic effect. Garnett and Baldwin, Cancer Research 46:2407-2412 (1986) have described the preparation of a drug-monoclonal antibody conjugate. Pastan et al., Cell 47:641-648 (1986) have reviewed the use of toxins conjugated to monoclonal antibodies for the therapy of various cancers. Goodwin and Meares, Cancer Supplement 80:2675-2680 (1997) have described the use of Yittrium-90 labeled monoclonal antibodies in various strategies to maximize the dose to tumor while limiting normal tissue toxicity. Other known cytotoxic radionuclides include Copper-67, Iodine-131, and Rhenium-186 all of which can be used to label monoclonal antibodies directed against the multimeric polypeptide antigen for the treatment of cancer of the breast.

The present invention will now be described by way of examples, which are meant to illustrate, but not to limit, the scope of the present invention.

EXAMPLES

Example 1: Identification of Breast Tissue Library Mammaglobin and BU101 Gene-Specific Clones

Library Comparison of Expressed Sequence Tags (EST's) or Transcript Images. Partial sequences of cDNA clone inserts, so-called "expressed sequence tags" (EST's), were derived from cDNA libraries made from breast tumor tissues, breast non-tumor tissues and numerous other tissues, both tumor and non-tumor and entered into a database (LIFESEQ™ database, available from Incyte Pharmaceuticals, Palo Alto, CA) as gene transcript images. See International Publication No. WO 95/20681. (A transcript image is a listing of the number of EST's for each of the represented genes in a given tissue library. EST's sharing regions of mutual sequence overlap are classified into clusters. A cluster is assigned a clone number from a representative 5' EST. Often, a cluster of interest can be extended by comparing its consensus sequence with sequences of other EST's which did not meet the criteria for automated

clustering. The alignment of all available clusters and single EST's represent a contig from which a consensus sequence is derived.) The transcript images then were evaluated to identify EST sequences that were representative primarily of the breast tissue libraries. These target clones then were ranked according to their abundance (occurrence) in the target libraries and their absence from background libraries. Higher abundance clones with low background occurrence were given higher study priority.

EST's corresponding to the consensus sequence of Mammaglobin (SEQUENCE ID NO 1) were found in 56.4% (22 of 39) of the breast tissue libraries. EST's corresponding to SEQUENCE ID NO 1, or fragments thereof were found in only 0.9% (7 of 754) of the other, non-breast libraries of the data base. Therefore, the EST's corresponding to SEQUENCE ID NO 1, or fragments thereof were found more than 60 times more often in breast than non-breast tissues.

Incorporated herein, by reference, are U.S. patent application Serial Number 08/697,106 filed on 8/19/96 which was abandoned in favor of continuation-in-part U.S. patent application Serial Number 08/912,149 filed on 8/15/97. The latter shows a set of contiguous and partially overlapping cDNA sequences and polypeptides encoded thereby, designated as mammaglobin and transcribed from breast tissue which are useful for the detecting, diagnosing, staging, monitoring, prognosticating, preventing or treating, or determining the predisposition of an individual to diseases and conditions of the breast such as breast cancer.

Similarly, EST's corresponding to the consensus sequence of BU101 (SEQUENCE ID NO 2) were found in 25.6% (10 of 39) of breast tissue libraries. EST's corresponding to SEQUENCE ID NO 2, or fragments or complements thereof, were found in only 1.1% (8 of 754) of the other, non-breast libraries of the database. Therefore, the EST's corresponding to SEQUENCE ID NO 2, or fragments or complements thereof, were found more than 24 times more often in breast than non-breast tissues.

Also incorporated, by reference, are U.S patent application Serial Number 08/697,105 filed on 8/19/96 which was abandoned in favor of continuation-in-part U.S. patent application Serial Number 08/912,276 filed on 8/15/97. The latter shows a set of contiguous and partially overlapping cDNA sequences and polypeptides encoded thereby, designated as BU101 and transcribed from breast tissue which are useful for the detecting, diagnosing, staging, monitoring, prognosticating, preventing

or treating, or determining the predisposition of an individual to diseases and conditions of the breast such as breast cancer.

When electronic Northern blots identifying the libraries in which EST's comprising the Mammaglobin (SEQUENCE ID NO 1) and BU101 (SEQUENCE ID NO 2) genes were compared, it appeared that Mammaglobin (SEQUENCE ID NO 1) and BU101 (SEQUENCE ID NO 2) were independently-expressed genes (Figures 1-2). The electronic Northerns revealed that some libraries contain Mammaglobin (SEQUENCE ID NO 1) but not BU101 (SEQUENCE ID NO 2) EST's while others contain BU101 (SEQUENCE ID NO 2) but not Mammaglobin (SEQUENCE ID NO 1) EST's.

In contrast to the lack of correlation between the mRNA expression of Mammaglobin (SEQUENCE ID NO 1) and BU101 (SEQUENCE ID NO 2) as found in the LifeSeq database, we observed a close correlation between their expression as determined in ribonuclease protection assays (Figure 3).

Analysis of the LIFESEQ™ database indicates a possible T/C polymorphism at position 254 in the BU101 nucleotide sequence (SEQUENCE ID NO 2). There were 33 occurrences of the C nucleotide variant (as in SEQUENCE ID NO 2) and eight occurrences of the T nucleotide variant in the database. The C nucleotide variant encodes a proline residue (CCG) whereas the T nucleotide variant (CTG) encodes a leucine residue at that position of the polypeptide.

Example 2: Production of Antibodies Against the Multimeric Polypeptide Complex

A. Production of Polyclonal Antisera.

1. Animal Immunization using Multimeric Polypeptide Complex as Immunogen. Female white New Zealand rabbits weighing 2 kg or more are used for raising polyclonal antiserum. One week prior to the first immunization, 5 to 10 ml of blood is obtained from the animal to serve as a non-immune prebleed sample.

Purified recombinant multimeric polypeptide complex (produced in accordance with example 7) is used to prepare the primary immunogen by emulsifying 0.5 ml of the protein complex at a concentration of 2 mg/ml in PBS (pH 7.2) with 0.5 ml of complete Freund's adjuvant (CFA) (Difco, Detroit, MI). The immunogen is injected into several sites of the animal via subcutaneous,

intraperitoneal, and/or intramuscular routes of administration. Four weeks following the primary immunization, a booster immunization is administered. The immunogen used for the booster immunization dose is prepared by emulsifying 0.5 ml of the same multimeric polypeptide complex used for the primary immunogen, except that the polypeptide now is diluted to 1 mg/ml with 0.5 ml of incomplete Freund's adjuvant (IFA) (Difco, Detroit, MI). Again, the booster dose is administered into several sites and can utilize subcutaneous, intraperitoneal and intramuscular types of injections. The animal is bled (5 ml) two weeks after the booster immunization and the serum is tested for immunoreactivity to the multimeric polypeptide complex, as described below. The booster and bleed schedule is repeated at 4 week intervals until an adequate titer is obtained. The titer or concentration of antiserum is determined by microtiter EIA as described in Example 3. An antibody titer of 1:500 or greater is considered an adequate titer for further use and study.

2. Animal Immunization using Peptide as Immunogen. Incorporated herein, by reference, are U.S. patent application Ser. No. 08/697, 105 filed on 8/19/96 which was abandoned in favor of U.S. patent application Ser. No. 08/912, 276 filed on 8/15/97, and U.S. patent application Ser. No. 08/697, 106 filed on 8/19/96 which was abandoned in favor of U.S. patent application Ser. No. 08/912, 149 filed on 8/15/97, which describes the production of antibodies against the individual polypeptide chains, including BU101 and Mammaglobin, of this multimeric polypeptide complex.

B. Production of Monoclonal Antibody.

1. Immunization Protocol Using Multimeric Polypeptide Complex as Immunogen. Mice are immunized using immunogens prepared as described in Example 7, except that the amount of the multimeric polypeptide complex for monoclonal antibody production in mice is one-tenth the amount used to produce polyclonal antisera in rabbits. Thus, the primary immunogen consists of 100 μg of the multimeric polypeptide complex in 0.1 ml of CFA emulsion; while the immunogen used for booster immunizations consists of 50 μg of the multimeric polypeptide complex in 0.1 ml of IFA. Hybridomas for the generation of monoclonal antibodies are prepared and screened using standard techniques. The methods used for monoclonal antibody development follow procedures known in the art such as those detailed in Kohler and Milstein, Nature 256:494 (1975) and reviewed in J.G.R. Hurrel, ed., Monoclonal Hybridoma Antibodies: Techniques and Applications, CRC

Press, Inc., Boca Raton, FL (1982). Another method of monoclonal antibody development which is based on the Kohler and Milstein method is that of L.T. Mimms et al., Virology 176:604-619 (1990), which is incorporated herein by reference.

The immunization regimen (per mouse) consists of a primary immunization with additional booster immunizations. The primary immunogen used for the primary immunization consists of 100 μg of the multimeric polypeptide complex in 50 μl of PBS (pH 7.2) previously emulsified in 50 μl of CFA. Booster immunizations performed at approximately two weeks and four weeks post primary immunization consist of 50 μg of the multimeric polypeptide complex in 50 μl of PBS (pH 7.2) emulsified with 50 μl IFA. A total of 100 μl of this immunogen is inoculated intraperitoneally and subcutaneously into each mouse. Individual mice are screened for immune response by microtiter plate enzyme immunoassay (EIA) as described in Example 3 approximately four weeks after the third immunization. Mice are inoculated either intravenously, intrasplenically or intraperitoneally with 50 μg of the multimeric polypeptide complex in PBS (pH 7.2) approximately fifteen weeks after the third immunization.

Three days after this intravenous boost, splenocytes are fused with, for example, Sp2/0-Ag14 myeloma cells (Milstein Laboratories, England) using the polyethylene glycol (PEG) method. The fusions are cultured in Dulbecco's Modified Eagle's Medium (DMEM) with the addition of L-glutamine, L-asparagine, L-arginine, folic acid, and containing 10% fetal calf serum (FCS), plus 1% hypoxanthine, aminopterin and thymidine (HAT). Bulk cultures are screened by microtiter plate EIA following the protocol in Example 3. Clones reactive with the multimeric polypeptide complex used as immunogen and non-reactive with other unrelated proteins are selected for final expansion. Clones thus selected are expanded, aliquoted and frozen in DMEM containing 10% FCS and 10% dimethyl sulfoxide.

2. Immunization Protocol Using Peptide as Immunogen. Mice were immunized using peptide/carrier immunogens (i.e., peptides conjugated to a carrier protein) prepared as described previously in U.S. patent application Ser. No. 08/697,105 filed on 8/19/96 which was abandoned in favor of U.S. patent application Ser. No. 08/912,276 filed on 8/15/97, and U.S. patent application Ser. No. 08/697,106 filed on 8/19/96 which was abandoned in favor of U.S. patent application Ser. No.

08/912, 149 filed on 8/15/97. The amount of the peptide/carrier protein immunogen used for monoclonal antibody production in mice was approximately one-tenth the amount used to produce polyclonal antisera in rabbits. Thus, the primary immunogen consisted of 100 μg of the peptide conjugated to a carrier protein in 0.1 ml of CFA emulsion; while the immunogen used for booster immunizations consisted of 50 μg of the peptide/carrier protein in 0.1 ml of IFA. Hybridomas for the generation of monoclonal antibodies were prepared and screened using standard techniques. The methods used for monoclonal antibody development followed procedures known in the art such as those detailed in Kohler and Milstein, Nature 256:494 (1975) and reviewed in J.G.R. Hurrel, ed., Monoclonal Hybridoma Antibodies: Techniques and Applications, CRC Press, Inc., Boca Raton, FL (1982). Another method of monoclonal antibody development which is based on the Kohler and Milstein method is that of L.T. Mimms et al., Virology 176:604-619 (1990), which is incorporated herein by reference.

The immunization regimen (per mouse) consisted of a primary immunization with additional booster immunizations. The primary immunogen used for the primary immunization consisted of 100 μg of the peptide/carrier protein complex in 50 μl of PBS (pH 7.2) previously emulsified in 50 μl of CFA. Booster immunizations performed at approximately two weeks and four weeks post primary immunization consisted of 50 μg of the peptide/carrier protein complex in 50 μl of PBS (pH 7.2) emulsified with 50 μl IFA. A total of 100 μl of this immunogen was inoculated intraperitoneally into each mouse. Individual mice were screened for immune response by microtiter plate enzyme immunoassay (EIA) as described in Example 3 approximately four weeks after the third immunization. Mice were inoculated intravenously with 25 μg of the peptide/carrier protein complex in PBS (pH 7.2) approximately fifteen weeks after the third immunization.

Three days after this intravenous boost, splenocytes were fused with Sp2/0-Ag14 myeloma cells (Milstein Laboratories, England) using the polyethylene glycol (PEG) method. The fusions were cultured in Dulbecco's Modified Eagle's Medium (DMEM) with the addition of L-glutamine, L-asparagine, L-arginine, folic acid, and containing 10% fetal calf serum (FCS), plus 1% hypoxanthine, aminopterin and thymidine (HAT). Bulk cultures were screened by microtiter plate EIA following the protocol in Example 3. Clones reactive with the multimeric polypeptide complex and the peptide used as immunogen and non-reactive with other unrelated proteins were

selected for final expansion. Supernatant from the final expansion was harvested and used for further characterisation. The hybridoma cells from the expansion growth were harvested, aliquoted, and frozen in DMEM containing 10% FCS and 10% dimethyl sulfoxide for storage.

3. Production of Ascites Fluid Containing Monoclonal Antibodies. Frozen hybridoma cells, prepared as described hereinabove, were thawed and placed into expansion culture. Viable hybridoma cells were inoculated intraperitoneally into Pristane treated mice. Ascitic fluid was removed from the mice, pooled, filtered through a 0.2 μ filter and subjected to an immunoglobulin class G (IgG) analysis to determine the volume of the Protein A column required for the purification.

4. Purification of Monoclonal Antibodies From Ascites Fluid or Cell Culture Supernatant. Monoclonal antibodies can be purified from ascites fluid or cell culture supernatant using a variety of methods including Protein A, Protein G, and Protein L column chromatography or precipitation. Monoclonal antibody H85C21 was purified using an Immunopure(G) IgG Purification kit (Pierce). Forty-seven milliliters of the H85C21 culture supernatant that contained 29 ug/mL of IgG was mixed with 100 mLs of the binding buffer (0.02 M sodium phosphate, pH 7.0) (Pierce). One hundred forty-five milliliters of the mixture were passed over a Protein G column that was equilibrated with the binding buffer. The column was then eluted with elution buffer (0.1 M glycine-HCl, pH 2.7) (Pierce). One milliliter fractions were collected in tubes that contained 100 uL of 1M sodium bicarbonate for neutralization. Fractions were monitored for absorbance at 280nm. Appropriate fractions were pooled and dialyzed against PBS (pH 7.2) overnight at 2-8°C. Absorbance at 280 nm indicated that 1.51 mg of IgG was recovered from the Protein G column. The purified monoclonal antibody thus prepared and characterized was placed at -80°C for long term storage.

5. Further Characterization of Monoclonal Antibody. The isotype and subtype of the monoclonal antibodies produced as described hereinabove was determined using an EIA microtiter plate assay. Briefly, the peptide immunogen was prepared at 1 mg/mL in 50 mM carbonate, pH 9.6 and 100 μl was placed in each well of an Immulon 2® High Binding microtiter plate (Dynex Technologies, Chantilly, VA). The plate was incubated for 14-18 hours at room temperature and then washed four times with deionized water. The wells were blocked by adding 200 μl of Superblock® (Pierce Chemical Company, Rockford, IL) to each well and incubated at room

temperature for 30 minutes before discarding the solution. Antisera obtained from immunized rabbits and mice, as described hereinabove, were diluted in a protein blocking agent (i.e., 3% Superblock® solution) in PBS containing 0.05% Tween-20 (monolaurate polyoxyethylene ether) (Sigma Chemical Company, St. Louis, MO) and 0.05% sodium azide and placed in each well of the coated microtiter plate. The wells were then incubated for one hour at room temperature. Each well was washed four times with deionized water. One hundred microliters (100 µl) of alkaline phosphatase-conjugated goat anti-mouse IgG (H + L), or IgG1, or IgG2a, or IgG2b, or IgG3 (Southern Biotech, Birmingham, AL), diluted 1:2000 in 3% Superblock® solution was added to each well. The wells were incubated for one hour at room temperature. Next, each well was washed four times with deionized water. One hundred microliters (100 µl) of para-nitrophenyl phosphate substrate (Kirkegaard and Perry Laboratories, Gaithersburg, MD) then was added to each well. The wells were incubated for thirty minutes at room temperature. The absorbance at 405 nm was read in each well. The results of the isotype testing are presented in Table 1.

Stability testing also can be performed on the monoclonal antibody by placing an aliquot of the monoclonal antibody in continuous storage at 2-8°C and assaying optical density (OD) readings throughout the course of a given period of time.

Table 1. Characterization of Monoclonal Antibodies

			Peptide	
Study #	Experiment #	Hybridoma	Immunogen	Isotype
390	1	H85C21	BU101.8	IgG1
390	1	H171C113	BU101.8	IgG1
390	2	H155C16	BU101.8	IgG1
394	4	H1C81	BU101.9	IgG1
394	4	H9C65	BU101.9	IgG1
394	4	H17C51	BU101.9	IgG1
394	4	H27C79	BU101.9	IgG1
394	4	H50C22	BU101.9	IgG1
394	4	H51C26	BU101.9	IgG1
394	4	H54C62	BU101.9	IgG1
394	4	H91C52	BU101.9	IgG1
394	4	H92C70	BU101.9	IgG1
394	4	H87C44	BU101.9	IgG1
394	5	H8C70	BU101.3	IgG1
394	5	H20C81	BU101.3	IgG1
394	5	H65C16	BU101.3	IgG1

394	5	H68C68	BU101.3	IgG1
394	5	H73C40	BU101.3	IgG1
394	5	H80C20	BU101.3	IgG1
394	5	H90C34	BU101.3	IgG1
394	5	H95C30	BU101.3	IgG1
392	17	H9C81	BU101.8	IgG1
392	17	H34C68	BU101.8	IgG1
392	9	H111C15	MAM.1	IgG1
392	9	H147C78	MAM.1	IgG1

Example 3: Enzyme Immunoassays

A. Microtiter Plate Direct Detection EIA.

The immunoreactivity of polyclonal and/or monoclonal antiserum (against either BU101 or Mammaglobin) toward the recombinant polypeptide complex (produced in accordance with Example 7 of the present application or Example 2 of U.S. patent application Ser. No. 09/215,818, filed on 12/18/98 (incorporated by reference)) was determined by means of a microtiter plate EIA.

For antibody titer measurements, pooled and dialysed recombinant polypeptide complex was prepared at 2 ug/mL in 50mM carbonate buffer, pH 9.6 and 100 µl was placed in each well of an Immulon 2® High Binding microtiter plate (Dynex Technologies, Chantilly, VA). For comparison, synthetic, full length BU101 polypeptide (SEQUENCE ID NO 6) and transiently transfected Mammaglobin M/H (as described hereinbelow in Example 7C) were prepared similarly. The plate was incubated for 14-18 hours at room temperature and then washed four times with deionized water. The wells were blocked by adding 200 µl of Superblock® (Pierce Chemical Company, Rockford, IL) to each well and the plates were incubated at room temperature for 30 minutes before discarding the solution. Antisera obtained from immunized rabbits and mice, as described hereinabove in Example 2, were diluted 1:10, 1:100, 1:1000, 1:10000, 1:100000 in a protein blocking agent (i.e., 3% Superblock® solution) in PBS containing 0.05% Tween-20 (monolaurate polyoxyethylene ether) (Sigma Chemical Company, St. Louis, MO) and 0.05% sodium azide and placed in each well of the coated microtiter plate. The wells were then incubated for one hour at room temperature. Each well was washed four times with deionized water. One hundred microliters (100 µl) of alkaline phosphatase-conjugated goat anti-mouse IgG (Southern Biotech, Birmingham, AL), diluted 1:2000

in 3% Superblock® solution was added to each well. The wells were incubated for one hour at room temperature. Next, each well was washed four times with deionized water. One hundred microliters (100 µl) of para-nitrophenyl phosphate substrate (Kirkegaard and Perry Laboratories, Gaithersburg, MD) then was added to each well. The wells were incubated for thirty minutes at room temperature. The absorbance at 405 nm was read in each well. The titer was designated as the dilution of antibody that resulted in an absorbance of 0.5 units at 405nm.

In addition to titers, apparent affinities [K_d(app)] were also determined for some of the anti-peptide antisera. In this case, pooled and dialysed recombinant polypeptide complex was prepared at dilutions of 1:3, 1:9, 1:27, 1:81, 1:243, 1:729, 1:2187, 1:6561, and 1:19683 in PBS and 100 µl was placed in each well of an Immulon 2® High Binding microtiter plate (Dynex Technologies, Chantilly, VA). For comparison, synthetic, full length BU101 polypeptide (SEQUENCE ID NO 6) and transiently transfected Mammaglobin M/H (as described hereinbelow in Example 7C) were prepared similarly. The plate was incubated for 14-18 hours at room temperature and then washed four times with deionized water. The wells were blocked by adding 200 µl of Superblock® (Pierce Chemical Company, Rockford, IL) to each well and the plates were incubated at room temperature for 30 minutes before discarding the solution. Antisera obtained from immunized rabbits and mice, as described hereinabove in Example 2, were diluted at an appropriate dilution in a protein blocking agent (i.e., 3% Superblock® solution) in PBS containing 0.05% Tween-20 (monolaurate polyoxyethylene ether) (Sigma Chemical Company, St. Louis, MO) and 0.05% sodium azide and placed in each well of the coated microtiter plate. The wells were then incubated for one hour at room temperature. Each well was washed four times with deionized water. One hundred microliters (100 µl) of alkaline phosphatase-conjugated goat anti-mouse IgG (Southern Biotech, Birmingham, AL), diluted 1:2000 in 3% Superblock® solution was added to each well. The wells were incubated for one hour at room temperature. Next, each well was washed four times with deionized water. One hundred microliters (100 µl) of para-nitrophenyl phosphate substrate (Kirkegaard and Perry Laboratories, Gaithersburg, MD) then was added to each well. The wells were incubated for thirty minutes at room temperature. The absorbance at 405 nm was read in each well. EIA microtiter plate assay results were used to derive the apparent dissociation constants [$K_{d(app)}$] based on an analog of the

Michaelis-Menten equation (V. Van Heyningen, Methods in Enzymology, Vol. 121, p. 472 (1986) and further described in X. Qiu, et al, Journal of Immunology, Vol. 156, p. 3350 (1996)):

$$[\text{Ag-Ab}] = [\text{Ag-Ab}]_{max} \times \frac{[\text{Ab}]}{[\text{Ab}] + K_d}$$

where [Ag-Ab] was the antigen-antibody complex concentration, $[\text{Ag-Ab}]_{max}$ was the maximum complex concentration, [Ab] was the antibody concentration, and K_d was the dissociation constant. During the curve fitting, the [Ag-Ab] was replaced with the background subtracted value of the OD_{405nm} at the given concentration of Ab. Both K_d, which corresponds to Kd(app), and $[OD_{405nm}]_{max}$, which corresponds to the $[\text{Ag-Ab}]_{max}$, were treated as fitted parameters. The software program Origin was used for the curve fitting. Figure 1 demonstrates the binding curves of three of the monoclonal antibodies that recognized the multimeric polypeptide complex. Kd(app) values are a measure of the relative affinities of the monoclonal antibodies to this complex. Kd(app) values are listed for each of the monoclonal antibodies in Table 2.

Table 2: Binding Properties of Monoclonal Antibodies

	Peptide	BU101.8[1]	Peptide	MamM/H[2]	RPC[3]	RPC[3]
Hybridoma	Immunogen	Titer	Epitope	Titer	Titer	Kd(app)
H85C21	BU101.8	800	BU101.5	ND	ND	128
H171C113	BU101.8	10	BU101.8	ND	ND	6
H155C16	BU101.8	1000	BU101.4	ND	ND	18
H1C81	BU101.9	8700	BU101.9/BU101.3	ND	190	27
H9C65	BU101.9	4600	BU101.9/BU101.3	ND	450	33
H17C51	BU101.9	2200	BU101.9/BU101.3	ND	100	22
H27C79	BU101.9	6100	BU101.9/BU101.3	ND	600	34
H50C22	BU101.9	4000	BU101.9/BU101.3	ND	50	21
H51C26	BU101.9	5900	BU101.9/BU101.3	ND	60	26
H54C62	BU101.9	900	BU101.9/BU101.3	ND	90	26
H91C52	BU101.9	6100	BU101.9/BU101.3	ND	250	25
H92C70	BU101.9	975	BU101.9/BU101.3	ND	90	35
H87C44	BU101.9	7500	BU101.9/BU101.3	ND	800	24
H8C70	BU101.3	5500	BU101.9/BU101.3	ND	400	26
H20C81	BU101.3	6000	BU101.9/BU101.3	ND	800	34
H65C16	BU101.3	8000	BU101.9/BU101.3	ND	500	24
H68C68	BU101.3	10,000	BU101.9/BU101.3	ND	500	31
H73C40	BU101.3	9000	BU101.9/BU101.3	ND	600	30
H80C20	BU101.3	5600	BU101.9/BU101.3	ND	100	36
H90C34	BU101.3	900	BU101.9/BU101.3	ND	65	27
H95C30	BU101.3	4900	BU101.9/BU101.3	ND	300	22
H9C81	BU101.8	300	BU101.8	ND	85	ND
H34C68	BU101.8	600	BU101.8	ND	150	ND

H111C15	MAM.1		MAM.1	50	100	76
H147C78	MAM.1		MAM.1	ND	10	ND

Legend:

1. BU101.8 is synthetic, full-length BU101
2. Mam M/H was prepared in accordance with Example 7C
3. RPC is recombinant polypeptide complex, produced in accordance with example 7C

B. Microtiter Plate Sandwich EIA.

Briefly, samples suspected of containing the multimeric polypeptide antigen are incubated in the presence of any combination of the following anti-BU101, anti-Mam, anti-α' polypeptide, anti-β' polypeptide, or anti-multimeric polypeptide antibody-coated microtiter wells in order to form antigen/antibody complexes. The microtiter wells then are washed and an indicator reagent comprising an antibody conjugated to a signal generating compound (i.e., enzymes such as alkaline phosphatase or horseradish peroxide) is added to the antigen/antibody complexes on the microtiter wells and incubated. The microtiter wells are washed and the bound antibody/antigen/antibody complexes are detected by adding a substrate (e.g., 4-methyl umbelliferyl phosphate (MUP), or OPD/peroxide, respectively), that reacts with the signal generating compound to generate a measurable signal. An elevated signal in the test sample, compared to the signal generated by a negative control, detects the presence of the multimeric polypeptide antigen. The presence of the multimeric polypeptide antigen in the test sample is indicative of a diagnosis of a breast disease or condition, such as breast cancer.

In an analogous manner, samples suspected of containing the multimeric polypeptide antigen are incubated in the presence of one or more of the following steroid (progesterone, aldosterone, androstenedione, corticosterone, cortisol, dehydroepiandrosterone, dihydrotestosterone, estradiol, estriol, estrone, hydroxyprogesterone, and testosterone) coated microtiter wells in order to form antigen/steroid complexes. The microtiter wells then are washed and an indicator reagent comprising either an antibody or a steroid conjugated to a signal generating compound is added to the steroid/antigen complexes on the microtiter wells and incubated. The microtiter wells are washed and the bound steroid/antigen/indicator reagent complexes are detected by adding a substrate that reacts with the signal generating compound to generate a measurable signal. An elevated signal in the test

sample, compared to the signal generated by a negative control, detects the presence of the multimeric polypeptide antigen. The presence of the multimeric polypeptide antigen in the test sample is indicative of a diagnosis of a breast disease or condition, such as breast cancer.

C. Microtiter Plate Competitive EIA.

The competitive binding assay uses a labeled polypeptide or protein complex that generates a measurable signal when the labeled polypeptide or protein complex is contacted with a microtiter well coated with an anti-polypeptide antibody. The labeled polypeptide is added to the multimeric polypeptide antibody-coated microtiter well in the presence of a test sample suspected of containing the multimeric polypeptide antigen, and incubated for a time and under conditions sufficient to form labeled peptide (or labeled protein)-bound antibody complexes and/or patient antigen-bound antibody complexes. The multimeric polypeptide antigen in the test sample competes with the labeled polypeptide (or protein) for binding sites on the microtiter well. The multimeric polypeptide antigen in the test sample results in a lowered binding of labeled peptide to the antibody-coated microtiter wells in the assay since antigen in the test sample and the peptide or protein compete for antibody binding sites. A lowered signal (compared to a control) indicates the presence of the multimeric polypeptide antigen in the test sample. The presence of the multimeric polypeptide antigen suggests the diagnosis of a breast disease or condition, such as breast cancer.

Similarly, the competitive binding assay uses a labeled polypeptide or protein complex that generates a measurable signal when the labeled polypeptide or protein complex is contacted with a microtiter well displaying a steroid. The labeled polypeptide is added to the steroid coated microtiter well in the presence of a test sample suspected of containing the multimeric polypeptide antigen, and incubated for a time and under conditions sufficient to form labeled peptide (or labeled protein)-bound steroid complexes and/or patient antigen-bound steroid complexes. The multimeric polypeptide antigen in the test sample competes with the labeled polypeptide (or protein) for binding sites on the microtiter well. The multimeric polypeptide antigen in the test sample results in a lowered binding of labeled peptide to the steroid-coated microtiter wells in the assay since antigen in the test sample and the peptide or protein compete for steroid binding sites. A lowered signal (compared

to a control) indicates the presence of the multimeric polypeptide antigen in the test sample. The presence of the multimeric polypeptide antigen suggests the diagnosis of a breast disease or condition, such as breast cancer.

The multimeric polypeptide complex which is provided and discussed hereinabove is useful as a marker of breast tissue disease, especially breast cancer. Tests based upon the appearance of this marker in a test sample such as tissue, blood, plasma or serum can provide low cost, non-invasive, diagnostic information to aid the physician to make a diagnosis of cancer, to help select a therapy protocol, or to monitor the success of a chosen therapy. This marker may appear in readily accessible body fluids such as blood, urine or stool as antigens derived from the diseased tissue which are detectable by immunological methods. This marker may be elevated in a disease state, altered in a disease state, or be a normal protein of the breast which appears in an inappropriate body compartment.

Example 4: Immunoprecipitation of the Multimeric Polypeptide Complex

Immune sera, obtained as described hereinabove in Example 2, is used to immunoprecipitate the multimeric polypeptide complex from solution prepared from tissue, blood, serum, or other bodily fluid. For tissue specimens, protein extracts are prepared by homogenizing tissue samples in 0.1M Tris-HCl (pH 7.5), 15% (w/v) glycerol, 0.2mM EDTA, 10 μg/ml leupeptin and 1.0 mM phenylmethylsulfonylfluoride [Kain et al., Biotechniques,17:982 (1994)]. Following homogenization, the homogenates are centrifuged at 2000 x g at 4°C for 5 minutes to separate supernatant from debris. Debris is re-extracted by homogenization with a buffer that is similar to above but also contains 0.1M Tricine and 0.1% SDS. Serum specimens can be used directly. Other bodily fluids may require preparation before immunoprecipitation.

The immunoprecipitation begins by coupling the antigen (if present) to the antibody by placing the sample (10-200 μl) in an Eppendorf tube. Bring the volume to 200 μl with dilution buffer (10mM tris-HCl, pH 8.0, 150mM NaCl, 0.1% Triton X-100, 0.025% sodium azide, 0.1% bovine serum albumin). Add the polyclonal serum (0.5 – 5 μl), hybridoma culture supernatant (10-100 μl), or ascites fluid (0.1 – 1 μl). Gently mix for 1.5 to 6 hours at room temperature. Precipitate the immune complex by adding 20 – 40 μl of 50% Protein A Sepharose slurry (Pharmacia Biotech, Piscataway, NJ). Gently mix for 1.5 to 16 hours. Centrifuge 1 minute at 200 x g.

Carefully remove supernatant and save pellet. Wash pellet with 1 mL of 10mM Tris-HCl, pH 8.0, 150mM NaCl, 0.025% sodium azide followed by 1 mL of 50mM Tris-HCl, pH 6.8. Again, carefully remove supernatant and save pellet. Dissociate the immune complex by adding 20-50 µl of SDS-PAGE buffer (50mM Tris-HCl, pH 6.8, 10% glycerol, 2% SDS, 2% beta-mercaptoethanol, 0.01% bromphenol blue) to the washed Protein A Sepharose beads. Cap and heat the sample at 100C for five minutes. Microfuge to pellet the Sepharose. Apply the supernatant to SDS-polyacrylamide gels and proceed with electrophoresis as described in Example 7. After SDS-PAGE, the antigen can be detected by protein staining, immunoblotting, and/or radiography.

Example 5: Purification of Serum Antibodies Which Specifically Bind to Polypeptides

Immune sera, obtained as described hereinabove in Example 2, is affinity purified using immobilized recombinant polypeptide complex prepared as in accordance with Example 7. An IgG fraction of the antiserum is obtained by passing the diluted, crude antiserum over a Protein A column (Affi-Gel protein A, Bio-Rad, Hercules, CA). Elution with a buffer (Binding Buffer, supplied by the manufacturer) removes substantially all proteins that are not immunoglobulins. Elution with 0.1M buffered glycine (pH 3) gives an immunoglobulin preparation that is substantially free of albumin and other serum proteins.

Immunoaffinity chromatography is performed to obtain a preparation with a higher fraction of specific antigen-binding antibody. The polypeptide used to raise the antiserum is immobilized on a chromatography resin, and the specific antibodies directed against its epitopes are adsorbed to the resin. After washing away non-binding components, the specific antibodies are eluted with 0.1 M glycine buffer, pH 2.3. Antibody fractions are immediately neutralized with 1.0 M Tris buffer (pH 8.0) to preserve immunoreactivity. The chromatography resin chosen depends on the reactive groups present in the polypeptide. If the polypeptide has an amino group, a resin such as Affi-Gel 10 or Affi-Gel 15 is used (Bio-Rad, Hercules, CA). If coupling through a carboxy group on the polypeptide is desired, Affi-Gel 102 can be used (Bio-Rad, Hercules, CA). If the polypeptide has a free sulfhydryl group, an organomercurial resin such as Affi-Gel 501 can be used (Bio-Rad, Hercules, CA).

Alternatively, spleens can be harvested and used in the production of hybridomas to produce monoclonal antibodies following routine methods known in the art as described hereinabove.

Example 6: Immunohistochemical Detection of the Multimeric Polypeptide Complex

Monoclonal antibodies, as described herein in Example 2, and listed in Table 3, were used to immunohistochemically label a formalin-fixed paraffin-embedded cell line (MB8, as described hereinbelow in Example 7) as well as malignant and normal breast tissues using standard procedures. D.L. Spector et al, In: Cells: A Laboratory Manual, Plainview, NY: Cold Spring Harbor Laboratory Press 1998. Briefly, 5μm sections were cut and placed on positively charged slides which were heated on a slide warmer at 60 °C for 30 minutes. The sections were rehydrated twice in zylene for 5 minutes each, twice in 100% ethanol for 1 minute each, three times in 95% ethanol for 1 minute each, and distilled water for 3 minutes. The sections were heated for 30 minutes in a Black and Decker Vegetable Steamer in 10 mM citrate buffer pH 6.0 and then cooled for 20 minutes to room temperature. The sections were washed in distilled water for 5 minutes and then blocked for 15 minutes with 1X casein (Dako Corp., Carpinteria, CA) diluted in Tris-buffered-saline [0.05 M Tris-HCl pH 7.6, 0.15 M NaCl] (TBS). The hybridoma culture supernatant was diluted 1:1 in TBS. This diluted supernatant was added to the sections and incubated at room temperature for 60 minutes, then washed twice in TBS for 5 minutes each. For detection, the LSAB+ kit (Dako Corp., Carpinteria, CA) was used. The sections were incubated with the link antibody for 30 minutes at room temperature and then washed twice in TBS for 5 minutes each. The sections then were incubated with streptavidin for 30 minutes at room temperature and washed twice in TBS for 5 minutes each. The sections were developed with BCIP/NBT/INT substrate (Dako Corp., Carpinteria, CA) for 15 minutes, placed in distilled water, and mounted with aqueous mounting media. The sections were viewed with a Nikon Optiphot II light microscope with a 10X objective and recorded with a Photometrics CoolSnap CCD camera and Metamorph version 4.0 software. Figures 2 and 3 illustrate the results of immunohistochemically staining two malignant breast sections, one normal breast section, and the MB8 cell line with monoclonal antibodies H9C65 and H95C30, respectively.

Table 3: Monoclonal antibodies that have been tested on formalin-fixed paraffin embedded MB8 cell line and tissue sections. All antibodies were tested on MB8 cell line. *Monoclonal antibodies that have been tested on two malignant breast tissues and one normal tissue. ¶Monoclonal antibodies that have been tested against ten malignant and three normal tissues.

Study #	Exp. #	Hybridoma
390	1	*H85C21
		H171C113sc16
390	2	*H155C16
392	4	*H1C81
		¶H9C65
		H17C51
		H27C79
		H50C22
		H51C26
		H54C62
		H87C44
		H91C52
		H92C70
392	5	H8C70
		H20C81
		H65C16
		H68C68
		*H73C40
		H80C20
		H90C34
		¶H95C30

Example 7: Stable Transfection and Expression of a Complex Comprising Mam M/H and BU101 M/H from Human Embryonic Kidney 293 Cells

A. Production of stable cell line, HEK293 - MB8

Incorporated by reference are U.S. patent application Ser. No.08/697,105, filed on 9/19/96 which was abandoned in favor of U.S. patent application Ser. No. 08/912,276 filed on 8/15/97, and U.S. patent application Ser. No. 08/697,106 filed on 8/19/96 which was abandoned in favor of U.S. patent application Ser. No. 08/912,149 filed on 8/15/97 which describe the production of BU101 myc/his (M/H) expression plasmids which utilized either clone 603148 or clone 2083578, and the Mam M/H expression plasmid which utilized clone 899895. These expression plasmids were retransformed into DH5 alpha cells, plated onto LB/ampicillin agar, and grown up in 10 ml of LB/ampicillin broth. The plasmids were purified using a QIAfilter™ Maxi

Kit (Qiagen, Chatsworth, CA) and were transfected into HEK293 cells [F.L. Graham et al., J. Gen. Vir. 36:59-72 (1977)].

The purified expression plasmids, as described supra, were transfected into HEK293 cells [F.L. Graham et al., J. Gen. Vir. 36:59-72 (1977)]. These cells are available from the American Type Culture Collection, 10801 University Boulevard, Manassas, Virginia 20110 under Accession No. CRL 1573. Transfection of the Mam M/H and BU101 M/H expression plasmids was performed using the cationic lipofectamine-mediated procedure described by P. Hawley-Nelson et al., Focus 15.73 (1993). Particularly, HEK293 cells were cultured in 10 ml DMEM media supplemented with 10% fetal bovine serum (FBS), L-glutamine (2 mM), sodium pyruvate (1 mM) and essential amino acids and freshly seeded into 60 mm culture plates at a density of 9 x 10^6 cells per plate. The cells were grown at 37 °C to a confluency of between 70% and 80% for transfection. Two micrograms of Mam M/H plasmid DNA and two micrograms of BU101 M/H plasmid DNA were added to 800 µl of unsupplemented DMEM medium (Gibco-BRL, Grand Island, NY). 8 ul of Plus Reagent (Gibco-BRL, Grand Island, NY) was added to this solution, which was then mixed briefly. 12 ul of Lipofectamine (LTI) was added to a second 800 µl portion of unsupplemented DMEM media. After a fifteen minute incubation, the two solutions were mixed and incubated at room temperature for an additional 15-30 minutes. During this time the culture medium was removed from the plates containing the HEK293 cells. The DMEM containing the Plus reagent:Lipofectamine:plasmid DNA complex was then overlaid onto the cells. The cells were incubated for 5 hr at 37°C and 5% CO_2, after which time, an additional 2 - 8 mL of DMEM with 20% FBS were added. After 18-24 hr, the old medium was aspirated, and the cells were overlaid with 5 mL of fresh DMEM with 5% FBS containing 400ug/ml G418, and the incubation was continued until 72 hrs had elapsed. Supernatants were analyzed for Mam M/H and BU101 M/H polypeptide expression by Western blot analysis.

At 72 hours post transfection, the cells were released from the dish by limited trypsinization and reseeded into 100 mm culture dishes in DMEM, 10%FBS, 400 ug/ml G418 at dilutions of 1:100, 1:1000 and 1:10000. These cultures were allowed to grow for 5-7 days, until well-isolated foci of cells were identified by microscopy. These foci were isolated by cloning cylinders, their cells released by limited trypsinization, and individual foci were transferred to separate wells in 24-well dishes, again in DMEM, 10% FBS, 400 ug/ml G418. After growth for 7-10 days, the

supernatants of each well were analysed for Mam M/H and BU101 M/H expression by Western blot analysis, as described hereinbelow. The clonal line labeled MB8 was found to express both Mam M/H and BU101 M/H in the supernatant. This line was expanded into 75 cm^2 flasks, and then passaged 1:30 three times, following which expression of Mam M/H and BU101 M/H was again reconfirmed to ensure stability of the insertion event. The final product of this procedure was a cell line derived from HEK293 cell line that expresses Mam M/H and BU101 M/H, which we have labeled MB8 (HEK293-MB8).

B. Analysis of Media

Aliquots of the supernatants from the MB8 cells were analyzed for the presence of both Mam M/H and BU101 M/H recombinant proteins. The aliquots were prepared in a reducing sample buffer (final concentration of 50 mM Tris pH 6.8, 10% glycerol (v/v), 2% sodium dodecyl sulfate (w/v), 2% beta-mercaptoethanol (v/v), 0.01% bromphenol blue (w/v)) and then electrophoresed on SDS-polyacrylamide gels (SDS-PAGE) using standard methods and reagents known in the art. (J. Sambrook et al., supra) Specifically, 40 µl of sample was added to 10 µl of sample buffer (5X) and the mixture was boiled for 5-10 minutes. Fifteen µl of that prepared sample was then loaded on a 10-20% Tricine gel, 1mm thick, (Novex, San Diego, CA) and electrophoresed at 110 V for approximately 90 minutes. These gels were then blotted overnight at 20V onto a solid medium such as nitrocellulose, and the Mam M/H and BU101 M/H protein bands were visualized using Western blotting techniques with a monoclonal antibody recognizing a myc epitope (Invitrogen, Carlsbad, CA) or polyclonal antisera recognizing either BU101 or Mammaglobin. Specifically, the nitrocellulose blot was removed and blocked with 0.2% I-block® (Tropix, Bedford, MA) for 60 minutes at room temperature. An appropriate amount of the primary antibody was then added. For example, the polyclonal antisera were used at a dilution of 1:5000 and the anti-myc epitope monoclonal antibody was used at a dilution of 1:5000. The primary antibody solution was exposed to the blot for 60 minutes at room temperature with shaking. The blot was then washed three times with I-block® solution. The secondary antibody, including either biotinylated goat anti-rabbit IgG or biotinylated goat anti-mouse IgG, was then prepared in I-block® solution at an appropriate dilution (1:5000) and exposed to the blot for 60 minutes at room

temperature with shaking. The blot was then washed three times with I-block® solution. The conjugate, alkaline phosphatase labeled streptavidin, was then prepared in I-block® solution at an appropriate dilution (1:10,000) and exposed to the blot for 30 minutes at room temperature with shaking. The blot was then washed three times with I-block® solution, followed by two times with assay buffer (20mM Tris (pH 9.8)/1 mM magnesium chloride). Twenty-five milligrams of 5-bromo-4-chloro-3-indolyl phosphate, BCIP, was dissolved in 0.5 mL of dimethylformamide. One hundred microliters of this BCIP solution was then mixed with 30 mLs of assay buffer and this BCIP substrate solution was exposed to the blot until bands were visible. The blot was then removed from the substrate solution and allowed to dry in the air. Electronic copies of the blots were obtained by scanning the dried blot.

Figure 4 contains three Western blots of supernatants harvested from the growth of the stably transfected clone, MB8. The fractionated proteins were detected under reducing conditions using a monoclonal antibody recognizing the myc epitope (Fig 4, blot 1), a polyclonal antisera recognizing BU101 (Fig 4, blot 2), or a polyclonal antisera recognizing Mam (Fig 4, blot 3). Lane 2 represents 35 mLs of supernatant harvested from a T75 flask. Lanes 3 and 4 represent 30 mLs of supernatant each harvested from a T150 flask. As can be seen, the Mam M/H and BU101 M/H protein concentration is approximately twice in lanes 3 and 4 as that observed in lane 2, which is consistent with the supernatant volume : growth area ratio.

Figure 4, blot 2 illustrates the Western blot, as described above, that was developed with a polyclonal antisera recognizing BU101. As shown, the media from the MB8 cell growth contains BU101 M/H, which is observed as two very close but discrete bands at approximately 11kD.

Figure 4, blot 3 illustrates the Western blot, as described above, that was developed with a polyclonal antisera recognizing Mam. As shown, the supernatant from the MB8 cell growth contains Mam M/H, which is observed as two species of approximately 20kD and 30kD in size. These species are consistent with those observed from earlier work (see U.S. patent application Ser. No. 09/215,218, filed on 12/18/98) involving the transient transfection of HEK293 cells with plasmids for both BU101 M/H and Mam M/H. The 20 kD and 30 kD bands are attributed to Mam M/H with one glycosylated Asn residue and Mam M/H with two glycosylated Asn

residues, respectively. It was noted that the predominant form of Mam M/H from the MB8 cell line was the more fully glycosylated form.

Figure 4, blot 1 illustrates the Western blot, as described above, that was developed with a monoclonal antibody recognizing a myc epitope. As shown, the media from the MB8 cell growth contains myc tagged Mam and BU101.

C. Nickel Chelation Chromatography

Supernatant from the growth of MB8 cells, as described supra, was applied to Chelating Sepharose Fast Flow (Pharmacia) charged with nickel for the purification of Mam M/H and BU101 M/H. Specifically, 40mLs of Chelating Sepharose Fast Flow was packed into a 16mm x 10cm column. Forty millilitres of nickel sulfate (0.1 M) in water were passed over the column to charge it with nickel. The column was washed and equilibrated with 10 mM sodium phosphate, 500 mM sodium chloride, pH 7.4. Two hundred twenty millilitres of supernatant from the growth of the MB8 cells were applied to the equilibrated column, and the histidine tagged proteins were eluted using a linear gradient of imidazole. The flow rate was 2 mLs/min; the gradient was 6.25 mM imidazole per milliliter of buffer; and the elution time was 80 minutes, creating an elution profile that went from 0 to 500 mM imidazole. Each 4 mL fraction was sampled. One hundred microliters of each fraction was applied to a well of a dot blot apparatus and the volume was suctioned through a piece of nitrocellulose. The nitrocellulose filter was then developed with the same procedure to develop Western blots, as described hereinabove in Example 7B, using a monoclonal antibody recognizing a myc epitope. Figure 5 (upper blot) illustrates the developed dot blot, which shows immunorecognition of material in fractions 20-35 by the anti-myc monoclonal antibody. These fractions correspond to elution conditions of 250 – 438 mM imidazole indicating the successful binding of the histidine tagged proteins to the nickel column and their elution with a histidine analogue. Fractions 20-35 were pooled and dialysed for a minimum of 4 hours each, against 2 x 4 L of phosphate buffered saline (PBS, 50mM phosphate, 150mM sodium chloride, pH 7.4) using Slide-a-Lysers (3500 MWCO). The pooled, dialysed, semi-purified MB8 supernatant was analysed for the presence of both Mam M/H and BU101 M/H recombinant proteins by Western blot.

For comparison, Mam M/H was similarly prepared. Supernatant containing Mam M/H was produced by the transient transfection of HEK293 cells with the Mam

M/H expression plasmid utilizing clone 899895. The methods and reagents for this procedure are described in U.S. patent application Ser. No. 08/697,106 filed on 8/19/96 which was abandoned in favor of U.S. patent application Ser. No. 08/912,149 filed on 8/15/97. One hundred twenty five milliliters of such supernatant was applied to a 40 mL, nickel charged, Chelating Sepharose Fast Flow column, as described supra. Column chromatography and analysis of fractions proceeded as described supra. Figure 5 (lower blot) illustrates the developed dot blot, which shows immunorecognition of material in fractions 19 – 36 by the anti-myc monoclonal antibody. Fractions 19 – 36 were pooled and dialysed for a minimum of 4 hours each, against 2 x 4 L of phosphate buffered saline using Slide-a-Lysers (3500 MWCO).

The two samples (semi-purified MB8 and Mam M/H supernatants) were prepared in a reducing sample buffer, electrophoresed, transferred to nitrocellulose, and developed with monoclonal and polyclonal antisera, as described hereinabove in Example 7B. In panels 1 – 8 of Figure 6, lane 1 represents the pooled, dialysed, semi-purified supernatant from the transient transfection of Mam M/H, and lane 2 represents the pooled, dialysed, semi-purified MB8 supernatant. Panels 1, 2, 3, and 4 were developed with polyclonal antisera recognizing either BU101 (10918, 10923, and 11543) or Mam (10931) respectively. Panels 1, 2, and 3 showed BU101 in the MB8 supernatant (lane 2) only. Panel 4 showed Mam in both the Mam M/H supernatant (lane 1) and in the MB8 supernatant (lane 2).

Panels 5, 6, 7, and 8 were developed with monoclonal antibodies recognizing either BU101 (H85, H68), Mam (H111), or myc, respectively. Again, the anti-BU101 monoclonals showed BU101 in the MB8 supernatant only (panels 5 and 6, lane 2). Panel 7 showed Mam in both Mam M/H supernatant (lane 1) and in the MB8 supernatant (lane 2). Panel 8 showed BU101 M/H in the MB8 supernatant (lane 2) and Mam M/H in both lanes.

A second set of Western blots were run on the same samples as described supra (MB8 supernatant and Mam M/H supernatant). In this experiment, the samples were prepared in a non-reducing sample buffer, electrophoresed, transferred to nitrocellulose, and developed with monoclonal and polyclonal antisera, as described supra. In panels 9 - 16, lane 1 represents the pooled, dialysed, semi-purified supernatant from the transient transfection of Mam M/H, and lane 2 represents the pooled, dialysed, semi-purified MB8 supernatant. Panels 9, 10, 11, and 12 were developed with polyclonal antisera recognizing either BU101 (10918, 10923, and

11543) or Mam (10931) respectively. Panels 9, 10, and 11, which were developed with anti-BU101 polyclonal antibodies identified two bands in the MB8 supernatant (lane 2) but nothing in the Mam M/H supernatant (lane 1). Panel 12, which was developed with an anti-Mam polyclonal antibody identified the same two bands in the MB8 supernatant (lane 2) which had been identified with the three anti-BU101 polyclonal antibodies. The anti-Mam antibody identified a number of bands in the Mam M/H supernatant (Panel 12, lane 1).

Panels 13, 14, 15 and 16 were developed with monoclonal antibodies recognizing either BU101 (H85, H68), Mam (H111), or myc, respectively. Again, the anti-BU101 monoclonal antibodies identified two bands in the MB8 supernatant (panels 13 and 14, lane 2). The anti-Mam monoclonal antibody identified the same two bands in the MB8 supernatant (panel 15, lane 2) as well as a number of bands in the Mam M/H supernatant (lane 1). The anti-myc monoclonal antibody also identified the same two bands in the MB8 supernatant (lane 2) as well as a number of bands in the Mam M/H supernatant (lane 1).

The MB8 supernatant contained two species that, under non-reducing conditions, were identified with anti-BU101 polyclonal and monoclonal antisera, anti-Mam polyclonal and monoclonal antisera, as well as the anti-myc monoclonal antibody. From panels 1 – 8, it was shown that these reagents (anti-BU101 and anti-Mam polyclonal and monoclonal antibodies) were not cross-reactive. Anti-BU101 reagents did not recognize Mam M/H and anti-Mam reagents did not recognize BU101 M/H. Therefore, the two bands detected in the MB8 supernatant by both anti-BU101 reagents and anti-Mam reagents contained both Mam M/H and BU101 M/H. The two bands are attributed to a complex comprising both BU101 M/H and Mam M/H, whereby Mam M/H may have different glycosylation states. Furthermore, this complex was dissociated by the addition of a reducing agent, such as beta-mercaptoethanol, to the non-reducing sample buffer for Western blots. Other reducing agents which may be utilized include, for example, dithiothreitol, beta-mercaptoethylamine, and tris(2-carboxyethyl)phosphine hydrochloride (TCEP).

D. Isoelectric Focusing

The isoelectric point of the proteins of interest was determined using the isoelectric function of the Wisconsin Sequence Analysis Package (Genetics Computer Group). The isoelectric point is a property of all proteins and is the pH at which the

protein has zero net charge. Thomas E. Creighton, ed., Proteins; Structures and Molecular Properties, 2nd edition, W H Freeman and Company, NY (1993). Table 4 lists the pI of the proteins of interest:

Table 4: Isoelectric Points of the Proteins of Interest

	pI	
Protein	no M/H tag	with M/H tag
Mammaglobin	3.8	4.6
BU101	8.4	7.6
Mam-BU101	4.4	5.9

The pooled, dialysed, semi-purified MB8 supernatant (from nickel chelation chromatography) and the pooled, dialysed, semi-purified Mam M/H supernatant (from nickel chelation chromatography) were prepared for isoelectric focusing using IEF 3-10 gels, cathode buffer pH 3-10, anode buffer, and sample buffer pH 3-10 (Novex) according to the manufacturer's instructions. Briefly, 50 uL of supernatant was added to 50 uL of sample buffer. Thirty microliters were loaded into each well of the gel. The gel was electrophoresed at 100 V for one hour, then 200 V for one hour, and finally 500 V for half an hour. The gel was blotted onto nitrocellulose overnight at 22 V constant voltage. The Western blot was developed as described hereinabove in Example 7B with anti-myc monoclonal antibody. Figure 7 shows the resulting Western blot. The Mam M/H supernatant was in lane 1 and the MB8 supernatant was in lane 2. The immunoreactive material was observed at the anode (bottom) end of the gel (pH 3), away from the cathode (top) end of the gel (pH 10). These results were consistent with the calculated pI of the proteins. The pI of Mam M/H was calculated to be 4.6 and the pI of Mam M/H and BU101 M/H together was calculated to be 5.9. Indeed, Mam M/H (lane 1) did focus at a lower pH than did the MB8 supernatant (lane 2).

E. Ion Exchange Chromatography

Semi-purified supernatant from the growth of MB8 cells, after nickel chelation chromatography, was further purified using anion exchange chromatography.

Specifically, 10 mLs of nickel purified MB8 supernatant, as described hereinabove in Example 7C, was dialysed against 2 L of 20 mM piperazine, pH 6.0. This material was applied to a Mono Q 5/5 column (Pharmacia) equilibrated with 20 mM piperazine, pH 6.0. The proteins were eluted using a linear gradient of sodium chloride. The flow rate was 1 mL/min; the gradient was 20 mM sodium chloride per milliliter of buffer; and the elution time was 50 minutes; creating an elution profile that went from 0 to 1000 mM sodium chloride. The sample loading time was 30 minutes prior to application of the elution gradient. Each 1 mL fraction was sampled. One hundred microliters of each fraction was boiled in the presence of 0.1% beta-mercaptoethanol, cooled, and was then applied to a well of a dot blot apparatus and the volume was suctioned through a piece of nitrocellulose. The nitrocellulose filter was then developed with the same procedure to develop Western blots, as described herein in Example 7B, using a monoclonal antibody recognizing a myc epitope. Figure 8 (upper blot) illustrates the developed dot blot, which shows immunorecognition of material in fractions 36-47 by the anti-myc monoclonal antibody. The immunoreactive material elutes between 120 - 340 mM sodium chloride, the center of the peak eluting at 180 mM sodium chloride. These results are consistent with the calculated isoelectric point of 5.9 for Mam M/H and BU101 M/H together, in that the material bound weakly to the anion exchange column at pH 6.0. The isoelectric point of BU101 M/H alone was calculated to be 7.6 and this protein would not be expected to bind to an anion exchange column at pH 6.0, since it would have a net positive charge. BU101 M/H would be expected to flow through the column during sample loading (the first 30 fractions). However, no anti-myc immunoreactive material was observed in these 30 flow-through fractions. These results are consistent with BU101 M/H having altered isoelectric properties due to covalent association with Mam M/H.

For comparison, Mam M/H supernatant was chromatographed on the same anion exchange column under the same conditions as described supra. Each 1 mL fraction was analysed as described supra. Figure 8 (lower blot) illustrates the developed dot blot, which shows immunorecognition of material in fractions 38 – 47 by the anti-myc monoclonal antibody. The immunoreactive material eluted between 160 – 340 mM sodium chloride. The elution profile on the dot blot indicated two peaks with centers at 200 mM sodium chloride (fraction 40) and 300 mM sodium chloride (fraction 45), unlike the material from the MB8 supernatant. These results

are consistent with the calculated isoelectric point of 4.6 for Mam M/H, in that the material bound moderately to the anion exchange column at pH 6.0.

F. Gel Filtration Chromatography

Semi-purified supernatant from the growth of MB8 cells, after nickel chelation chromatography, was further purified using gel filtration chromatography. Specifically, 4.65 mLs of nickel purified MB8 supernatant, as described supra, was concentrated to 650 uL and 200 uL of the concentrate was applied to a 10mm x 30cm column of Superose 12 (Pharmacia). The column was run with a single buffer of PBS (50mM phosphate, 150mM sodium chloride, pH 7.4) at a flowrate of 0.4 mL/min. The column was calibrated with molecular weight standards available from Pharmacia. The resulting standard curve for molecular weight determination is illustrated in Figure 9, which demonstrates the relationship between molecular weight and elution volume.

The elution of the myc-his tagged Mam and BU101 proteins from the Superose 12 column was monitored by immunorecognition with anti-myc monoclonal antibody. Each 0.4 mL fraction was sampled. One hundred microliters of each fraction was boiled in the presence of 0.1% beta-mercaptoethanol, cooled, and was then applied to a well of a dot blot apparatus and the volume was suctioned through a piece of nitrocellulose. The nitrocellulose filter was then developed with the same procedure to develop Western blots, as described hereinabove in Example 7, using a monoclonal antibody recognizing a myc epitope. Figure 10 illustrates the developed dot blot, which shows immunorecognition of material in fractions 33-36 by the anti-myc monoclonal antibody. Fractions 33, 34, 35, and 36 have elution volumes of 13.2 mL, 13.6 mL, 14.0 mL and 14.4 mL, respectively. The center of this peak (13.8 mL) corresponds to a molecular weight of 56kD. These results are consistent with an association between Mam M/H and BU101 M/H. As was shown in Figure 6, BU101 M/H and Mam M/H, as individual species, have vastly different molecular weights (~11 kD and ~30 kD). Figure 9 showed the performance of the Superose 12 column, and the capability of it to separate such individual species. Instead, the elution profile showed a single peak with an average molecular weight of 56kD, consistent with the species identified in the MB8 supernatant under non-reducing conditions (Figure 6, panels 9-16).

Example 8: Identification of the Multimeric Polypeptide Complex from Human Tissue

A. Preparation of Tissue Extract

Two tenths of a gram of breast cancer tissue, which had been snap frozen and stored at –70C was prepared for Western blot analysis. Protein extracts were prepared by homogenizing tissue samples in 0.1 M tris-HCl (pH 7.5), 15% (w/v) glycerol, 0.2 mM EDTA, 10 ug/mL leupeptin, and 1.0 mM phenylmethylsulfonylfluoride. S. R. Kain et al., Biotechniques 17: 982 (1994) Following homogenization, the homogenates were centrifuged at 4C for 5 minutes to separate supernatant from debris. For protein quantitation, 2-5 uL of supernate was added to 1.5mL of Coomassie Protein Reagent (Pierce Chemical Co., Rockford, IL) and the absorbance was read at 595nm.

B. Analysis of Tissue Extract

The breast cancer tissue extract, as described supra, was analysed by Western blot. Other samples included recombinant Mam M/H and BU101 M/H and another breast cancer tissue extract, as described hereinbelow. The samples were prepared in both a reducing sample buffer (final concentration of 50 mM Tris pH 6.8, 10% glycerol (v/v), 2% sodium dodecyl sulfate (w/v), 2% beta-mercaptoethanol (v/v), 0.01% bromphenol blue (w/v)) and a non-reducing sample buffer (final concentration of 50mM Tris pH 6.8, 10% glycerol (v/v), 2% sodium dodecyl sulfate (w/v), and 0.01% bromphenol blue (w/v)) and then electrophoresed on SDS-polyacrylamide gels (SDS-PAGE) using standard methods and reagents known in the art. (J. Sambrook et al., supra) Specifically, 40 µl of sample was added to 10 µl of sample buffer (5X) and the mixture was boiled for 5-10 minutes. Fifteen µl of that prepared sample was then loaded on a 10-20% Tricine gel, 1mm thick, (Novex, San Diego, CA) and electrophoresed at 110 V for approximately 90 minutes. These gels were then blotted overnight at 20V onto a solid medium such as nitrocellulose, and the Mam and BU101 protein bands were visualized using Western blotting techniques with monoclonal or polyclonal antisera recognizing either BU101 or Mammaglobin. Specifically, the nitrocellulose blot was removed and blocked with 0.2% I-block® (Tropix, Bedford, MA) for 60 minutes at room temperature. An appropriate amount of

the primary antibody was then added. For example, the polyclonal antisera was used at a dilution of 1:5000 and the monoclonal culture supernatant was used at a dilution of 1:50. The primary antibody solution was exposed to the blot for 60 minutes at room temperature with shaking. The blot was then washed three times with I-block® solution. The secondary antibody, including either biotinylated goat anti-rabbit IgG or biotinylated goat anti-mouse IgG, was then prepared in I-block® solution at an appropriate dilution (1:5000) and exposed to the blot for 60 minutes at room temperature with shaking. The blot was then washed three times with I-block® solution. The conjugate, alkaline phosphatase labeled streptavidin, was then prepared in I-block® solution at an appropriate dilution (1:10,000) and exposed to the blot for 30 minutes at room temperature with shaking. The blot was then washed three times with I-block® solution, followed by two times with assay buffer (20mM Tris (pH 9.8)/1 mM magnesium chloride). Twenty-five milligrams of 5-bromo-4-chloro-3-indolyl phosphate, BCIP, was dissolved in 0.5 mL of dimethylformamide. One hundred microliters of this BCIP solution was then mixed with 30 mLs of assay buffer and this BCIP substrate solution was exposed to the blot until bands were visible. The blot was then removed from the substrate solution and allowed to dry in the air. Electronic copies of the blots were obtained by scanning the dried blot.

Figure 11 illustrates the Western blots of the breast cancer tissue extracts and the recombinant myc-his tagged Mam and BU101 proteins. The upper blot was developed with a monoclonal antibody recognizing BU101 and the lower blot was developed with a polyclonal antisera recognizing Mam. Four samples were analysed on the two blots. Samples in lanes 4-7 were prepared with a reducing sample buffer (as described hereinabove in Example 7C) and samples in lanes 10-13 were prepared with a non-reducing sample buffer. The first sample was a tissue specimen from a patient with invasive, poorly differentiated adenocarcinoma (breast) which is represented in lanes 4 and 10. The second sample was the transiently expressed, myc-his tagged, Mam and BU101 protein complex which is represented in lanes 5 and 11. The third sample was the stably expressed, myc-his tagged, Mam and BU101 protein complex from MB8 cells which is represented in lanes 6 and 12. The fourth sample was a tissue specimen from another patient with adenocarcinoma (breast) which is represented in lanes 7 and 13.

The upper blot was developed with a monoclonal antibody against BU101. In lanes 4 – 7, the samples have been reduced and in lanes 10 – 13, the samples have not been reduced. Under reducing conditions, BU101 was observed in the tissue specimens of lanes 4 and 7, and recombinant BU101 M/H was observed in lanes 5 and 6 (weakly). The myc-his tag added approximately 3 kD to the molecular weight of BU101 and this effect was manifested as a slightly higher band on the Western blot. BU101 was detected as a single species in all 4 samples, although some background staining was observed. Under non-reducing conditions (lanes 10 – 13), BU101 was observed at a much higher molecular weight than when the samples were reduced. The tissue specimens of lanes 10 and 13 showed one predominant species (indicated by a line in Figure 11) when probed with the anti-BU101 monoclonal antibody. In contrast, the recombinant BU101 M/H of lanes 11 and 12 showed multiple species (indicated by lines in Figure 11). The observation of multiple species with the myc-his tagged proteins is attributed to the multiple glycosylation states of mammaglobin in this recombinant system which were observed in the lower blot of Figure 11.

The lower blot was developed with a polyclonal antisera against Mammaglobin. In lanes 4 – 7, the samples were reduced and in lanes 10 – 13, the samples were not reduced. Under reducing conditions, Mam was observed in the tissue specimens of lanes 4 and 7, and recombinant Mam M/H was observed in lanes 5 and 6. The myc-his tag added approximately 3 kD to the molecular weight of Mam and this effect was manifested as slightly higher bands on the Western blot. Mam was detected as a single species in the tissue specimens but as multiple species in the recombinant systems. These multiple bands in the recombinant systems are attributed to the multiple glycosylation states of the Mam protein (as described in U.S. patent application Ser. No. 09/215,818 (filed on 12/18/98 and incorporated by reference). Briefly, Mam contains two Asn residues which can be glycosylated. The 3 bands are attributed to Mam with both Asn residues glycosylated, or one Asn residue glycosylated, or no Asn residues glycosylated. It was noted that the stable cell line (MB8, lane 6) produced more fully glycosylated Mam than did the transient expression system (lane 5). Under non-reducing conditions (lanes 10 – 13), Mam was observed at a slightly higher molecular weight than when the samples were reduced. The tissue specimens of lanes 10 and 13 showed one predominant species (indicated by a line in Figure 11) when probed with the anti-Mam polyclonal antibody. In

contrast, the recombinant Mam M/H of lanes 11 and 12 showed multiple species (indicated by lines in Figure 11).

By direct comparison of the upper and lower blots, it was observed that the same band was detected in the breast cancer tissue specimens by both anti-BU101 and anti-Mam antibodies under non-reducing conditions (lanes 10 and 13). These results are consistent with an association between BU101 and Mam which can be removed with a reducing agent. Such an association might be disulfide linkage between the two species.

C. Ion Exchange Chromatography

The breast cancer tissue extract derived from the patient with invasive, poorly differentiated adenocarcinoma (breast) (Figure 11, lanes 4 and 10) was purified using anion exchange chromatography. Specifically, the 800 uL extract was applied to a Mono Q 5/5 column (Pharmacia) equilibrated with 20 mM piperazine, pH 6.0. The proteins were eluted using a linear gradient of sodium chloride. The flow rate was 1 mL/min; the gradient was 20 mM sodium chloride per milliliter of buffer; and the elution time was 50 minutes; creating an elution profile that went from 0 to 1000 mM sodium chloride. The sample loading time was 30 minutes prior to application of the elution gradient. Each 1 mL fraction was sampled. One hundred microliters of each fraction was boiled in the presence of 0.1% beta-mercaptoethanol and was then applied to a well of a dot blot apparatus and the volume was suctioned through a piece of nitrocellulose. The nitrocellulose filter was then developed with the same procedure to develop Western blots, as described hereinabove in Example 7B, using polyclonal antisera recognizing either BU101 or Mam. Figure 12 illustrates the developed dot blots, which show immunorecognition of material in fractions 41-48 by both anti-Mammaglobin polyclonal antisera and anti-BU101 polyclonal antisera. Fractions 41 – 48 correspond to a salt concentration of 220 – 360 mM sodium chloride. These results are consistent with an isoelectric point less than 4.6. Table 5 summarizes the elution conditions from the anion exchange column with the corresponding pI of the eluted protein or protein complex.

Table 5: Chromatographic Properties at pH 6.0

Protein	pI	Fraction Number	Salt Concentration

Mam M/H + BU101 M/H	5.9	36-47	120-340mM
Mam M/H	4.6	38-47	160-340
Multimeric Polypeptide Complex (Tissue Extract)	?	41-48	220-360

The isoelectric point of BU101 (no M/H tag) was calculated to be 8.4 (Table 4). The isoelectric point of Mam (no M/H tag) was calculated to be 3.8 (Table 4). The isoelectric point of a complex comprising one BU101 polypeptide and one Mam polypeptide was calculated to be 4.4, a value consistent with the chromatographic properties observed. Furthermore, a complex comprising one BU101 polypeptide and one Mam polypeptide is consistent with the dot blot analysis of the fractions, where both anti-BU101 polyclonal antisera and anti-Mam polyclonal antisera recognized fractions 41 – 48.

To analyze the material further, each positive fraction (fractions 41-48) was run on a Western blot under both reducing and non-reducing conditions. Two identical blots were produced for development with either anti-Mam polyclonal antisera or anti-BU101 polyclonal antisera. In Figure 13, the upper blot was developed with anti-BU101 polyclonal antisera, and the lower blot was developed with anti-Mam polyclonal antisera.

In the upper blot, which was developed with anti-BU101 polyclonal antisera, the fractions were run both in the reduced state (left side of the blot) and the non-reduced state (right side of the blot). In the reduced state, BU101 was observed in each fraction as a single band of low molecular weight (~ 7 kD). Under non-reducing conditions, a single, broad band was detected at a higher molecular weight (~ 23 – 34 kD).

In the lower blot, which was developed with anti-Mam polyclonal antisera, the fractions were run both in the reduced state (left side of the blot) and the non-reduced state (right side of the blot). In the reduced state, Mam was observed in each fraction as a single, broad band of molecular weight 17 – 23 kD. Under non-reducing conditions, a single, broad band was detected at a higher molecular weight (~ 23 – 34 kD). Comparison of the two blots demonstrated that anti-BU101 and anti-Mam polyclonal antisera had detected the same band under non-reducing conditions. The molecular weight of the non-reduced species correlated with the sum of one BU101

polypeptide with one Mam polypeptide. These data are consistent with a complex comprising Mam and BU101, which are disulfide linked.

D. Gel Filtration Chromatography

Semi-purified breast cancer tissue extract, as described supra, was further purified using gel filtration chromatography. Specifically, fractions 41 – 48 from ion exchange chromatography, as described supra, were concentrated to 100 uL using a Centriplus 30 concentrator spun at 3000 x g for 25 minutes. The retentate was applied to a 10mm x 30cm column of Superose 12 (Pharmacia). The column was run with a single buffer of PBS (50mM phosphate, 150mM sodium chloride, pH 7.4) at a flowrate of 0.4 mL/min. The column was calibrated with molecular weight standards available from Pharmacia. The resulting standard curve for molecular weight determination is illustrated in Figure 9, which demonstrates the relationship between molecular weight and elution volume.

The elution of Mam and BU101 from the Superose 12 column was monitored by immunorecognition with both anti-BU101 and anti-Mam polyclonal antisera. Each 0.4 mL fraction was sampled. One hundred microliters of each fraction was boiled in the presence of 0.1% beta-mercaptoethanol and was then applied to a well of a dot blot apparatus and the volume was suctioned through a piece of nitrocellulose. Two dot blots were prepared in this manner. The nitrocellulose filters were then developed with the same procedure to develop Western blots, as described hereinabove in Example 7B, using anti-BU101 polyclonal antisera on one nitrocellulose filter and anti-Mam polyclonal antisera on the other nitrocellulose filter. Figure 14 illustrates these developed dot blots, which show immunorecognition of material in fractions 34-39 by both the anti-Mam and anti-BU101 polyclonal antisera. Fractions 34, 35, 36, 37, 38, and 39 have elution volumes of 13.6, 14.0, 14.4, 14.8, 15.2, and 15.6 mLs respectively. The molecular weight range for these fractions corresponds to ~63 kD - ~20 kD. The center of this peak corresponds to a molecular weight of ~40 kD.

The detection of BU101 in the same fractions as Mam from gel filtration chromatography is consistent with a complex comprising both BU101 and Mam. As was shown in Figure 13, BU101 and Mam, as individual species, have vastly different molecular weights (~7 kD and ~23 kD). Figure 9 showed the performance of the Superose 12 column, and the capability of it to separate such individual species. Instead, the elution profile showed a single peak with an average molecular weight of

40kD, consistent with the species identified in the breast tissue extract under non-reducing conditions (Figure 13). This data does not support the presence of BU101 and Mam as isolated polypeptides. Furthermore, the identification of a complex comprising both BU101 and Mam from breast cancer tissue indicates its biological relevance in human tissue and disease.

Example 9: Enhanced Immunorecognition of a Complex Comprising Mam M/H and BU101 M/H using a Pretreatment Protocol

Semi-purified supernatants from both the growth of HEK293-MB8 cells and the transient transfection of Mam M/H (after nickel chelation chromatography as described in Example 7C) were subjected to different pretreatment protocols including non-ionic detergent, anionic detergent, or reducing agent, with or without heat. Specifically, aliquots of Tween 20, sodium dodecyl sulfate (SDS), or beta-mercaptoethanol (β-ME) were added to 100 uL of semi-purified supernatant, as described supra, for a final concentration in additive of 0.1%, 0.25%, 0.5%, 1.0%, 2.0%, and 4.0%. A duplicate set of samples was treated with heat for 5 minutes and cooled. The treated samples were then loaded into a dot blot apparatus and the volume was suctioned through a piece of nitrocellulose. The nitrocellulose filter was then developed with the same procedure to develop Western blots, as described hereinabove in Example 7B, using an anti-myc monoclonal antibody.

Pre-treatment of the semi-purified supernatant containing Mam M/H produced similar results to those produced in the pre-treatment of semi-purified supernatant from HEK293-MB8 cells (Figure 15). Immunorecognition of the polypeptides with respect to the anti-myc monoclonal antibody was enhanced with some pretreatment protocols. Tween 20, with or without heat treatment, had no effect on the immunorecognition of the anti-myc monoclonal antibody to the myc-his tagged polypeptides. SDS treatment improved the immunorecognition of the anti-myc monoclonal antibody to the myc-his tagged polypeptides, at the low concentrations of 0.1% and 0.25% better than at higher concentrations (> 0.5%). This effect was observed both in the presence and absence of heat. Beta-mercaptoethanol treatment improved the immunorecognition of the anti-myc monoclonal antibody to the myc-his tagged polypeptides at all concentrations and the effect was observed both in the

presence and absence of heat. Thus, immunorecognition of the polypeptides was shown to be enhanced with some pretreatment protocols.

Example 10: Enhanced Breast CancerPanel

A. Background

Genes that are expressed in highly tissue or disease-specific manner provide possible targets for therapeutics, early detection of cancer, and monitoring of disease burden during and after treatment. Further, genes of this type that code for secreted or shed proteins may allow for serum detection of the product facilitating our ability to specifically detect the cancer in all circumstances.

To this end, the inventors worked towards identification and characterization of such genes that are specifically expressed in breast epithelium. Two new genes that are highly restricted in their expression to the breast epithelium emerged from a directed screen of the Incyte LifeSeq Database: one of these is a newly discovered uteroglobin termed BU101, while the other is a novel gene with mucin-like properties termed BS106. The expression of both of these genes is largely limited to normal and neoplastic breast epithelium.

In the current study, the inventors have measured the expression of these two markers in comparison with two other candidate markers for detecting breast epithelium, mammaglobin and cytokeratin 19 (CK). Mammaglobin has been proposed as a breast-specific gene while cytokeratin has been used extensively to detect epithelium in axillary lymph nodes, bone marrow, and in peripheral blood. Using both semi-quantitative end-point PCR and quantitative real-time PCR, we compared the expression of these four genes in a series of primary breast cancers and uninvolved lymph nodes.

As anticipated, CK was highly sensitive in detecting all breast cancers; however, amplification of specific products was also seen from lymph nodes from non-cancer patients. The three breast markers, mammaglobin, BU101, and BS106 were more specific than CK; however, each of these markers also failed to detect a small partially overlapping subset of breast cancers. Therefore, while no one of these markers efficiently detects *all* breast cancers, a combination of two or more may achieve a very high sensitivity in assaying for circulating or occult breast cancer cells.

The development of these markers for the detection of tumor cells in the bone marrow and in circulation will be further discussed.

B. Marker Expression in Primary Cancers:

The inventors measured expression of three breast specific markers (BU101, Mammaglobin, and BS106) in a series of 101 primary breast cancers and 17 lymph nodes from patients without breast cancer using quantitative RT-PCR. All cancers were positive for cytokeratin 19 by this same assay.

Expression of a representative group of 20 cancers is shown in Figure 19 A-C. For each marker, a small fraction of cancers had undetectable levels of mRNA by this assay. However, all cancers contained detectable levels of at least one marker. This indicates the likely necessity of combining two or more markers in any assay. Three of the normal lymph nodes had detectable CK19 expression.

Note that the scales for "Relative Expression" vary significantly among the three markers. Wide variations in the level of expression were noted for each of the markers. Target mRNA quantitation for all three markers was spread over 5 orders of magnitude (100,000 fold). All PCR reactions were run on agarose gels to verify the presence of a band of the correct size. Even though we could measure mRNA levels through a range of 10^5, authentic PCR products were detected in all samples with non-zero values by qRT-PCR.

C. General Assay Methods

Equal amounts of total RNA from each cancer was isolated and converted into cDNA by random priming. Primers specific for each marker were used in PCR amplifications on an ABI 7700 cycler with the addition of Sybr Green to each reaction. Expression values were derived using the built-in crossing-point analysis software and normalized to the expression level of each marker in the breast cancer cell line, SKBR-3.

Protein extracted from breast tissue was fractionated using a MonoQ column at pH6 and eluted using increasing salt gradient. Fractions were assayed for BU101 and mammaglobin using polyclonal antisera raised against specific peptides. Signal for both proteins was seen in fractions 42-48 (Figure 20 A & B). These fractions were then run on SDS-PAGE, under reducing or non-reducing conditions. Gels were

blotted and probed with the same polyclonal antibodies as in (Figure 20 A & B). The MamBu complex is detected using both antibodies under non-reducing conditions (Figure 20 C & D). Further studies indicate that when these proteins are expressed in the same cell, the MamBu complex is the predominant form that is secreted.

Significant associations were observed between high BS106 expression and increased HER2/neu IHC staining. Also, as previously described, cancers expressing high levels of BU101 typically co-express mammaglobin at high levels as well.There is a weaker correlation between ER/PR negativity and high levels of mammaglobin.

In conclusion the inventors have identified two new breast-specific expression markers – BU101 and BS106. BU101 is a member of the uteroglobin family and is found in a secreted complex with mammaglobin. BU101 and mammaglobin are typically co-expressed in breast cancers. BS106 is a novel gene encoding a mucin-like protein and demonstrates breast tissue specificity in expression.

Preliminary studies suggest BS106 expression correlates with HER2/neu expression. Expression levels of all three markers are independent of tumor size, stage, and hormone receptor status. Expression of one or more of these markers is present in virtually all primary breast cancers. However, not all cancers express each marker indicating that a combination of two or more of these markers will be needed to detect all breast cancers.